STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
WHM vs placebo														
Liu 2021	Kumar 2019	UC	Curcumin	Clinical response	4 weeks	NR	NR	28/25	17/28	13/25	NR	NR	*NR*	High
				Clinical remission		NR	NR		NR	NR	NR	NR	*NR*	High
	Sugimoto 2019	UC	Curcumin	Clinical response	12 weeks	NR	NR	NR	NR	NR	NR	NR	*NR*	High
				Clinical remission		NR	NR		NR	NR	NR	NR	*NR*	High
	Masoodi 2018	UC	Curcumin	Clinical response	4 weeks	NR	NR	28/28	NR	NR	NR	NR	*NR*	Low
				Clinical remission		NR	NR		NR	NR	NR	NR	*NR*	Low
	Banerjee 2017	UC	Curcumin	Clinical response	12 weeks	NR	NR	22/25	12/22	5/25	NR	NR	*NR*	Low
				Clinical remission		NR	NR		NR	NR	NR	NR	*NR*	Low
				Endoscopic remission		NR	NR		5/22	0/25	NR	NR	*NR*	Low
	Kedia 2017	UC	Curcumin	Clinical response	8 weeks	NR	NR	29/33	6/29	12/33	NR	NR	*NR*	Low
				Clinical remission		NR	NR		9/29	9/33	NR	NR	*NR*	Low
	Lang 2015	UC	Curcumin	Clinical response	4 weeks	NR	NR	26/24	17/26	3/24	NR	NR	*NR*	Low
				Clinical remission		NR	NR		14/26	0/24	NR	NR	*NR*	Low
	Singla 2014	UC	Curcumin	Clinical response	8 weeks	NR	NR	23/22	13/23	8/22	NR	NR	*NR*	Low
				Clinical remission		NR	NR		10/23	5/22	NR	NR	*NR*	Low
	Hanai 2006	UC	Curcumin	Clinical response	24 weeks	NR	NR	45/44	NR	NR	NR	NR	*NR*	Low
				Clinical remission		NR	NR		41/45	31/44	NR	NR	*NR*	Low
	Rastegarpanah 2015	UC	St Mary's thistle	Clinical response	24 weeks	Symptoms	NR	42/38	NR	NR	NR	NR	*NR*	Some concerns
				Clinical remission		NR	NR		NR	NR	NR	NR	*NR*	Some concerns
	Dryden 2013	UC	EGCG (green tea extract)	Clinical response	8 weeks	NR	NR	15/4	10/15	0/4	NR	NR	*NR*	Some concerns
				Clinical remission		NR	NR		8/15	0/4	NR	NR	*NR*	Some concerns

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
	Footnotes:	Authors also report PP results. Data not extracted here.												
Chandan 2020	Masoodi 2018	UC	Curcumin	Clinical response	4 weeks	proportion with improvement in SCCAI		28/28	16/28	8/28	NR	NR	*NR*	Low
				Clinical remission		NR	NR		NR	NR	NR	NR	*NR*	Low
	Banerjee 2017	UC	Curcumin	Clinical response	12 weeks	Mayo score	decrease ≥ 3	22/25	12/22	5/25	NR	NR	*NR*	Some concerns
				Clinical remission		NR	NR		NR	NR	NR	NR	*NR*	Some concerns
				Endoscopic remission		Histology	score ≤ 1		5/22	0/25	NR	NR	*NR*	Low
	Kedia 2017	UC	Curcumin	Clinical response	8 weeks	UCDAI	decrease ≥ 3	29/33	6/29	12/33	NR	NR	*NR*	Low
				Clinical remission		UCDAI	score ≤ 2		9/29	9/33	NR	NR	*NR*	Low
	Lang 2015	UC	Curcumin	Clinical response	4 weeks	UCDAI	decrease ≥ 3	26/24	17/26	3/24	NR	NR	*NR*	Low
				Clinical remission		UCDAI	score <3		14/26	0/24	NR	NR	*NR*	Low
	Singla 2014	UC	Curcumin	Clinical response	8 weeks	UCDAI	decrease ≥ 3	23/22	13/23	8/22	NR	NR	*NR*	Low
				Clinical remission		UCDAI	score <3		10/23	5/22	NR	NR	*NR*	Low
	Shivakumar 2011	UC	Curcumin	Clinical response	8 weeks	NR	NR	18/18	NR	NR	NR	NR	*NR*	Low
				Clinical remission		fecal calprotectin	improvement		15/18	9/18	NR	NR	*NR*	Low
	Hanai 2006	UC	Curcumin	Clinical response	24 weeks	NR	NR	45/44	NR	NR	NR	NR	*NR*	Low
				Clinical remission		CAI	score ≤ 4		43/45	36/44	NR	NR	*NR*	Low
	Footnotes:													
Coelho 2020*	Sadeghi 2019	UC	Curcumin	Clinical response	8 weeks	NR	NR	35/35	NR	NR	NR	NR	*NR*	Low
				Clinical remission		SCCAI score	proportion with remission		83.90%	43.80%	NR	<0.05	*Favours intervention*	Low
				HRQoL		IBDQ-9	higher is better		NR	NR	NR	<0.05	*Favours intervention*	Low
				Fecal urgency		proportion with reduced urgency			60%	28.60%	NR	<0.05	*Favours intervention*	Low

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Coelho 2020* cont'd	Masoodi 2018	UC	Curcumin	General wellbeing	4 weeks	proportion with improvement		28/28	64%	39.30%	NR	<0.05	*Favours intervention*	Low
				Clinical response		SCCAI score	higher is worse		1.71 (NR)	2.68 (NR)	NR	<0.05	*Favours intervention*	Low
				Clinical remission		NR	NR		NR	NR	NR	NR	*NR*	Low
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Goulart 2020*	Sadeghi 2019	UC	Curcumin	Clinical response	8 weeks	SCCAI score	proportion with decrease ≥3	35/35	30/35	18/35	RD 0.34 (0.014, 0.54)	NR	*Favours intervention*	Low
				Clinical remission		SCCAI score	proportion with score ≤2		26/35	14/35	RD 0.34 (0.13, 0.56)	NR	*Favours intervention*	Low
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Zheng 2020	Footnotes:	Authors report PP results. Data not extracted here.												
Grammatikopoulou 2018*	Hanai 2006	UC	Curcumin	Clinical response	24 weeks	DAI	end of treatment score	45/44	1.0 (2.0)	2.2 (2.3)	NR	NR	*NR*	Low
				Endoscopic remission		Endoscopic index	end of treatment score		0.8 (0.6)	1.6 (1.6)	NR	NR	*NR*	Low
	Lang 2015	UC	Curcumin	Clinical response	4 weeks	DAI	end of treatment score	26/24	NR	NR	NR	NR	*NR*	Low
				Clinical remission		Endoscopic index	end of treatment score		1.35 (1.19)	2.25 (0.88)	NR	NR	*NR*	Low
	Kedia 2017	UC	Curcumin	Clinical response	8 weeks	UCDAI	end of treatment score	29/33	3.4 (3.1)	3.8 (2.8)	NR	NR	*NR*	Low
				Clinical remission		Endoscopic index	end of treatment score		NR	NR	NR	NR	*NR*	Low

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
	Banerjee 2017	UC	Curcumin	Clinical response	12 weeks	DAI	end of treatment score	22/25	NR	NR	NR	NR	*NR*	Some concerns
				Clinical remission		Endoscopic index	end of treatment score		NR	NR	NR	NR	*NR*	Some concerns
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Kafil 2017	Madisch 2007	Collagenous colitis	Boswellia	Clinical remission	6 weeks	stool frequency < 3 per day	% with improvement	16/15	7/16 (43.8%)	4/15 (26.7%)	NR	0.25	*No difference*	Low
	Footnotes:													
Kim 2017*	Rastegarpanah 2015	UC	St Mary's thistle	Clinical remission	24 weeks	DAI failure to maintain remission	% with remission failure	42/38	7/42	17/38	RR 0.37 (0.17, 0.80)	NR	*Favours intervention*	High
	Lang 2015	UC	Curcumin	Clinical remission	4 weeks	failure to achieve remission	% with remission failure	26/24	12/26	24/24	RR 0.47 (0.31, 0.71)	NR	*Favours intervention*	Some concerns
	Dryden 2013	UC	EGCG (green tea extract)	Clinical remission	8 weeks	failure to achieve remission	% with remission failure	16/4	8/16	4/4	RR 0.56 (0.32, 0.97)	NR	*Favours intervention*	High
	Sandbom 2013	UC	Andrographis extract	Clinical remission	8 weeks	failure to achieve remission	% with remission failure	149/75	96/149	56/75	RR0.86 (0.72, 1.03)	NR	*No difference*	Some concerns
	Hanai 2006	UC	Curcumin	Clinical remission	24 weeks	failure to maintain remission	% with remission failure	45/44	2/45	8/44	RR 0.24 (0.05, 1.09)	NR	*No difference*	Low
	Langmead 2004	UC	Aloe vera gel	Clinical remission	4 weeks	failure to achieve remission	% with remission	30/14	21/30	13/14	RR 0.75 (0.57, 0.99)	NR	*Favours intervention*	Low
	Hallert 1991**	UC	Psyllium husk	Clinical response	8 weeks	self-reported symptoms	% with improvement	36 (NR/NR)	64%	24%	NR	<0.001	*Favours intervention*	Some concerns

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Kim 2017* cont'd	Holtmeier 2011	Crohn's disease	Boswellia serrata resin extract	Clinical remission	52 weeks	failure to maintain remission	% with remission failure	42/40	19/42	19/40	RR 0.95 (0.60, 1.52)	NR	*No difference*	Low
	Sandbom 2010	Crohn's disease	Andrographis extract	Clinical remission	8 weeks	failure to achieve remission	% with remission failure	51/50	36/51	43/50	RR 0.82 (0.67, 1.01)	NR	*No difference*	Some concerns
	Omer 2007	Crohn's disease	Wormwood	Clinical remission	10 weeks	failure to achieve remission	% with remission failure	20/20	7/20	20/20	RR 0.37 (0.21, 0.65)	NR	*Favours intervention*	Some concerns
	Footnotes:	*Kim 2017 reports the proportion of participants who *fail to achieve (or maintain)* remission. The data are inverted in our evidence synthesis to correlate with other reviews that report the proportion of participants who *achieve* or maintain remission.												
Langhorst 2015*	Langmead 2004	UC	Aloe vera gel	Disease activity	4 weeks	SCCAI score	higher is worse	30/14	NR	NR	NR	<0.05	*Favours intervention*	Low
				HRQoL		IBDQ	higher is better		NR	NR	NR	NR	*Favours comparator*	Low
	Holtmeier 2011	Crohn's disease	Boswellia serrata resin extract	Disease activity	52 weeks	CDAI	higher is worse	42/40	NR	NR	NR	>0.05	*No difference*	Low
				HRQoL		IBDQ	higher is better		NR	NR	NR	>0.05	*No difference*	Low
	Omer 2007	Crohn's disease	Wormwood	Disease activity	10 weeks	CDAI	higher is worse	20/20	NR	NR	NR	<0.05	*Favours intervention*	Some concerns
				HRQoL		IBDQ	higher is better		NR	NR	NR	>0.05	*No difference*	Some concerns
				Emotional wellbeing		HAM-D	higher is worse		NR	NR	NR	>0.05	*No difference*	Some concerns
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
WHM vs inactive control														
Kim 2017*	Fernández-Bañares 1999	UC	Psyllium seeds	Clinical remission	52 weeks	failure to maintain remission	% with remission failure	69 (NR/NR)	NR	NR	RR 0.85 (0.42, 1.72)	0.65	*No difference*	High
	Krebs 2012	Crohn's disease	Wormwood	Clinical remission	6 weeks	failure to achieve remission	% with remission failure	20 (NR/NR)	NR	NR	RR 0.25 (0.07, 0.90)	NR	*Favours intervention*	High

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
	Footnotes:	*Kim 2017 reports the proportion of participants who *fail to achieve (or maintain)* remission. The data are inverted in our evidence synthesis to correlate with other reviews that report the proportion of participants who *achieve (* or maintain) remission.												
Langhorst 2015*	Krebs 2012	Crohn's disease	Wormwood	Disease activity	6 weeks	CDAI	higher is worse	20 (NR/NR)	NR	NR	NR	<0.05	*Favours intervention*	High
				HRQoL		IBDQ	higher is better		NR	NR	NR	<0.05	*Favours intervention*	High
				Emotional wellbeing		HAM-D	higher is worse		NR	NR	NR	<0.05	*Favours intervention*	High
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
WHM vs active control														
Langhorst 2015*	Tang 2011	UC	Andrographis extract vs mesalazine	Clinical remission	8 weeks	NR	NR	120 (60/60	improvement over baseline	improvement over baseline	NR	NR	*Noninferior*	Low
	Langhorst 2013	UC	Combination vs mesalazine	Clinical remission	6 weeks	NR	NR	97 (48/49)	NR	NR	NR	NR	*Noninferior*	Low
				Clinical response		NR	NR		NR	NR	NR	NR	*Noninferior*	Low
	Gerhardt 2001	Crohn's disease	Boswellia extract vs mesalazine	Symptom severity	8 weeks	CAI	higher is worse	102 (50/52)	NR	NR	NR	NR	*No difference*	Low
				Clinical remission		NR	NR		NR	NR	NR	NR	*No difference*	Low
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Rahimi 2013*	Tang 2011	UC	Andrographis extract vs mesalazine	Clinical remission	8 weeks	% with 100% improvement		120 (60/60	11/53	9/55	NR	NR	*Not reported*	Low
				Clinical response		% with some(>25%) improvement		120 (60/60	40/53	45/55	NR	NR	*Not reported*	Low
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Ernst 2008	Gerhardt 2001	Crohn's disease	Boswellia extract vs mesalazine	Symptom severity	8 weeks	CAI	higher is worse	102 (50/52)	192 (114)	163 (96)	NR	NR	*Noninferior*	Low
	Footnotes:	*Only RCT data not included elsewhere are extracted here												

Abbreviations: C, Comparator; Confidence interval; CAI, Crohn's activity index; DAI, disease activity index; IBDQ-9, inflammatory blwoel disease questionnaire; I, intervention; NR, not reported; RoB, risk of bias; SCCAI, simple clinical colitis activity index; SD, standard deviation; UC, ulcerative colitis

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
WHM vs placebo														
Black 2020*	**8 RCTs**	IBS	Peppermint oil	Global improvement	4 to 12 weeks	Failure to achieve improvement in global IBS symptoms		**442/382**	NR	NR	**RR 0.63 (0.48, 0.83)**	NR	*Favours intervention*	Low
	7 RCTs		Ispaghula husk					**246/253**	NR	NR	**RR 0.78 (0.59, 1.02)**	NR	*No difference*	Low
	8 RCTs	IBS	Peppermint oil	Abdominal pain	4 to 12 weeks	Failure to achieve improvement		**442/382**	NR	NR	**RR 0.66 (0.48, 0.91)**	NR	*Favours intervention*	Low
	7 RCTs		Ispaghula husk					**246/253**	NR	NR	**RR 0.88 (0.61, 1.27)**	NR	*No difference*	Low
	Footnotes:	*The authors conduct and report network meta-analysis results. Individual study data are not reported.												
Hawrelak 2020	Davis 2006	IBS	Aloe vera juice	Global improvement	4 weeks	% with changes in IBS-SSS		31/27	11/31	6/27	RR 1.60 (0.68, 3.74)	NR	*Favours intervention*	Some concerns
	Hutchings 2011	IBS	Aloe vera juice	HRQoL	20 weeks	IBS quality of life		110	NR	NR	NR	NR	*No difference*	Some concerns
				Global improvement		GI symptom rating scale			NR	NR	NR	NR	*No difference*	Some concerns
	Storsrud 2015	IBS	Aloe vera juice	Global improvement	4 weeks	% with changes in IBS-SSS		68 (33/35)	18/33	11/35	RR 1.74 (0.97, 3.10)	0.09	*Favours intervention*	Low
				Abdominal pain		IBS-SSS subscale pain severity			NR	NR	NR	0.011	*Favours intervention*	Low
				Emotional functioning		Hospital Anixety and Depression Scale			NR	NR	NR	NR	*Not reported*	Low
				Bowel transit time		Bowel transit time			NR	NR	NR	NR	*Not reported*	Low
	Rees 1979	IBS	Peppermint oil	Defecation frequency	3 weeks	changes in Stool frequency		18	NR	NR	NR	NR	*No difference*	Low
				Global improvement		% with changes in IBS symptom score			13/18	5/18	RR 2.6 (1.17, 5.78)	< 0.005	*Favours intervention*	Low
	Evans 1982	IBS	Peppermint oil	Global improvement	2 weeks	Change in symptom score		20	NR	NR	NR	< 0.005	*Favours intervention*	Some concerns
				Global improvement		% with changes in daily symptom score			24/29	5/29	RR 4.8 (2.13, 10.04)	< 0.01	*Favours intervention*	Low

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Hawrelak 2020 cont'd	Dew 1984	IBS	Peppermint oil	Abdominal pain or	2 weeks	% with Adequate relief		29/29	NR	NR	NR	<0.001	*Not reported*	Low
				Defecation frequency		changes in Stool frequency			NR	NR	NR	NR	*Not reported*	Low
	Nash 1986	IBS	Peppermint oil	Global improvement	2 weeks	% with changes in IBS symptom score		41/41	13/41	17/41	RR 0.76 (0.43, 1.36)	NR	*Favours comparator*	Some concerns
				Abdominal pain		% with improvement			NR	NR	NR	NR	*No difference*	Some concerns
				Defecation frequency		changes in Stool frequency			NR	NR	NR	NR	*No difference*	Some concerns
	Lawson 1988	IBS	Peppermint oil	Abdominal pain	4 weeks	% with improvement		25	NR	NR	NR	NR	*No difference*	Low
				Bloating		% with improvement			NR	NR	NR	NR	*No difference*	Low
				Defecation frequency		increase in Stool frequency			NR	NR	NR	< 0.05	*Favours intervention*	Low
	Lech 1988	IBS	Peppermint oil	Global improvement	4 weeks	% with changes in IBS symptom score		23/24	13/23	6/24	RR 2.26 (1.04, 4.93)	0.02	*Favours intervention*	Some concerns
				Abdominal pain		% with improvement			12/23	6/24	RR 2.09 (0.94, 4.63)	0.1	*No difference*	Some concerns
				Defecation frequency		changes in defecation pattern			NR	NR	NR	NR	*No difference*	Some concerns
	Weiss 1988	IBS	Peppermint oil	Global improvement	3 weeks	% with changes in IBS symptom score		30/30	17/30	4/30	RR 4.25 (1.62, 11.15)	< 0.001	*Favours intervention*	Some concerns
	Wildgrube 1988	IBS	Peppermint oil	Abdominal pain	2 weeks	% with improvement		40	NR	NR	NR	< 0.05	*Favours intervention*	Some concerns
				Bloating		% with improvement			NR	NR	NR	< 0.05	*Favours intervention*	Some concerns
	Carling 1989	IBS	Peppermint oil	Global improvement	2 weeks	Changes in IBS symptoms score		30/14	NR	NR	NR	0.063	*No difference*	Some concerns
				Global improvement		% with improvement			17/30	5/14	RR 1.59 (0.74, 3.42)	0.001	*Favours intervention*	Some concerns

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Hawrelak 2020 cont'd	Schneider 1990	IBS	Peppermint	Abdominal pain	6 weeks	Changes in abdominal symptoms		60	NR	NR	NR	0.02	*Favours intervention*	Some concerns
				Global improvement		Changes in IBS symptoms score			NR	NR	NR	0.002	*Favours intervention*	Some concerns
				Defecation frequency		changes in stool frequency			NR	NR	NR	NR	*No difference*	Some concerns
	Liu 1997	IBS	Peppermint oil	Abdominal pain	4 weeks	% with improvement in abdominal pain		55/55	41/55	21/55	RR 1.95 (1.35, 2.83)	< 0.05	*Favours intervention*	High
				Defecation frequency		% with improvement			NR	NR	NR	< 0.05	*Favours intervention*	High
	Kline 2001	IBS	Peppermint oil	Global improvement	2 weeks	% with change in IBS symptoms score		25/25	15/21	9/21	RR 1.67 (0.95, 293)	< 0.002	*Favours intervention*	High
				Abdominal pain		% with improvement in abdominal pain			16/25	4/25	RR 4.0 (1.55, 10.29)	< 0.001	*Favours intervention*	High
	Capanni 2005	IBS	Peppermint oil	Global improvement	12 weeks	% with change in overall symptoms		91/87	73/91	31/87	RR 2.25 (1.87, 3.04)	< 0.02	*Favours intervention*	Some concerns
	Cappello 2007	IBS	Peppermint oil	Global improvement	4 weeks	% with change in IBS symptoms score		28/29	18/28	10/29	RR 1.88 (1.05, 3.31)	0.05	*Favours intervention*	Some concerns
				Abdominal pain		% with improvement in abdominal pain			NR	NR	NR	0.05	*Favours intervention*	Some concerns
	Merat 2010	IBS	Peppermint oil	Abdominal pain	8 weeks	% free from abdominal pain		90	NR	NR	NR	< 0.001	*Favours intervention*	Some concerns
				HRQoL		SF-36 bodily pain, general health, social functioning,			NR	NR	NR	< 0.05	*Favours intervention*	Some concerns
	Alam 2013	IBS	Peppermint oil	Abdominal pain	6 weeks	Improvement in pain scores		74	NR	NR	NR	< 0.05	*Favours intervention*	Some concerns
	Cash 2016	IBS	Peppermint oil	Global improvement	4 weeks	Changes in IBS symptoms score		72	mean 40% reduction	mean 24.3% reduction	NR	0.0246	*Favours intervention*	Some concerns
	Mosaffa-Jahromi 2016	IBS	Peppermint oil	Abdominal pain	4 weeks	% with improvement in abdominal pain		40/40	NR	NR	NR	< 0.001	*Favours intervention*	Low
				Bloating		% with improvement in abdominal pain			NR	NR	NR	0.004	*Favours intervention*	Low
				Global improvement		Changes in IBS symptoms			21/40	14/40	RR 1.5 (0.90, 2.51)	NR	*Favours intervention*	Low

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Hawrelak 2020 cont'd	Pedersen 1998	IBS	Appital	Global improvement	8 weeks	Changes in IBS symptoms score		59	NR	NR	NR	0.081	*No difference*	Some concerns
	Madisch 2004	IBS	Iberogast	Abdominal pain	4 weeks	Abdominal pain scale		208	NR	NR	NR	0.0009	*Favours intervention*	Low
				Global improvement		Changes in IBS symptoms score			NR	NR	NR	0.001	*Favours intervention*	Low
			Iberogast-II	Abdominal pain		Abdominal pain scale			NR	NR	NR	0.0005	*Favours intervention*	Low
				Global improvement		Changes in IBS symptoms score			NR	NR	NR	0.003	*Favours intervention*	Low
	Brinkhaus 2005	IBS	Curcumin	Global improvement	18 weeks	Changes in IBS symptoms score		106	NR	NR	NR	NR	*No difference*	Some concerns
				Emotional functioning		psycological stress caused by IBS			NR	NR	NR	NR	*No difference*	Some concerns
				Abdominal pain		Changes in IBS related pain (VAS)			NR	NR	NR	NR	*No difference*	Some concerns
	Vejdani 2006	IBS	Lemon balm+ Mentha spicata+ coriander	Abdominal pain	8 weeks	Changes in abdominal pain		32	NR	NR	NR	0.016	*Favours intervention*	High
				Bloating		Changes in bloating severity			NR	NR	NR	0.002	*Favours intervention*	High
	Saito 2010	IBS	St. John's wort	Global improvement	12 weeks	Changes in bowel symptom score		70	NR	NR	NR	NR	*Favours comparator*	Low
				Adequate relief		% with adequate relief			NR	NR	NR	NR	*Favours comparator*	Low
	Bortolotti 2011	IBS	Capsicum	Abdominal pain	6 weeks	Changes in abdominal pain		50	NR	NR	NR	NR	*No difference*	High
				Defecation frequency		Changes in defecation frequency			NR	NR	NR	NR	*No difference*	High
				Bloating		Changes in bloating			NR	NR	NR	NR	*No difference*	High

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Hawrelak 2020 cont'd	Tilburg 2014	IBS	Ginger	Adequate relief	4 weeks	Changes in Adequate Relief rating scale		45	NR	NR	NR	NR	*No difference*	Some concerns
				Global improvement		Changes in IBS severity Scale			NR	NR	NR	NR	*No difference*	Some concerns
	Brown 2015	IBS	Horse chestnut + peppermint + Schinopsis	Bloating	2 weeks	Changes in bloating score (7-point Likert scale)		16	NR	NR	NR	< 0.001	*Favours intervention*	Some concerns
				Constipation		Changes in constipation score (7-point Likert scale)			NR	NR	NR	0.0034	*Favours intervention*	Some concerns
	Footnotes													
Tan 2020	Cappello 2007	IBS	Peppermint oil	Global improvement	4 weeks	Symptom score improvement		28/29	18/28	10/29	RR 1.86 (1.05, 3.31)	NR	*Favours intervention*	Some concerns
	Davis 2006	IBS	Aloe vera	Global improvement	4 weeks	Symptom score improvement		31/27	11/31	6/27	RR 1.6 (0.68, 3.74)	NR	*Favours intervention*	Some concerns
	Liu 1997	IBS	Peppermint oil	Abdominal pain	4 weeks	Improvement in abdominal pain		52/49	41/52	21/49	RR 1.84 (1.29, 2.62)	NR	*Favours intervention*	Some concerns
	Mosaffa-Jahromi 2016	IBS	Anise oil	Global improvement	4 weeks	IBS symptom resolution		40/40	30/40	14/40	RR 2.14 (1.35, 3.39)	NR	*Favours intervention*	Low
			Peppermint oil			IBS symptom resolution		40/40	21/40	14/40	NR	NR	*Not reported*	Low
	Tilburg 2014	IBS	Ginger	Global improvement	4 weeks	Symptom score improvement		30/15	12/30	9/15	RR 0.67 (0.37, 1.22)	NR	*Favours comparator*	Some concerns
	Portincasa 2016	IBS	Curcumin + fennel oil	Global improvement	4 weeks	Complete symptom free rate		60/61	15/60	4/61	RR 3.81 (1.34, 10.83)	NR	*Favours intervention*	Some concerns
				HRQoL		IBS quality of life			NR	NR	NR	0.003	*Favours intervention*	Some concerns
	Vejdani 2006	IBS	Lemon balm+ Mentha spicata+ coriander	Global improvement	8 weeks	Overall relief of IBS symptoms		14/18	8/14	3/18	RR 3.43 (1.11, 10.59)	NR	*Favours intervention*	Some concerns
				Abdominal pain		Relief of symptoms			NR	NR	NR	NR	*Not reported*	Some concerns
				Bloating		Relief of symptoms			NR	NR	NR	NR	*Not reported*	Some concerns

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Tan 2020 cont'd	Merat 2010	IBS	Peppermint oil	Abdominal pain	8 weeks	Free from pain/discomfort		45/45	14/45	6/45	RR 2.33 (0.98, 5.53)	NR	*Favours intervention*	Some concerns
				HRQoL		SF-36			NR	NR	NR	NR	*Not reported*	Some concerns
	Storsrud 2015	IBS	Aloe vera	Global improvement	4 weeks	Symptom score improvement (≥50 point reduction)		33/35	18/33	11/35	RR 1.74 (0.97, 3.10)	NR	*Favours intervention*	Low
				Emotional functioning		Hamilton Anxiety Depression Scale			NR	NR	NR	NR	*No difference*	Low
	Saito 2010	IBS	St. John's wort	Adequate relief	12 weeks	Adequate relief of symptoms		35/35	11/35	21/35	RR 0.52 (0.3, 0.92)	NR	*Favours comparator*	High
				HRQoL		IBS quality of life			NR	NR	NR	0.7	*No difference*	High
	Footnotes:													
Alammar 2019	Cash 2016	IBS	Peppermint oil	Global improvement	4 weeks	Improvement in global symptoms in IBS		34/37	13/34	7/37	RR 2.02 (0.92, 4.46)	NR	*Favours intervention*	High
				Abdominal pain		Proportion of patients with improvement			14/34	8/37	RR 1.9 (0.91, 3.96)	NR	*Favours intervention*	High
	Capanni 2005	IBS	Peppermint oil	Global improvement	12 weeks	Improvement in global symptoms in IBS		91/87	73/91	31/87	RR 2.25 (1.67,3.04)	NR	*Favours intervention*	Some concerns
				Abdominal pain		Proportion of patients with improvement			34/91	16/87	RR 2.03 (1.21,3.41)	NR	*Favours intervention*	Some concerns
	Cappello 2007	IBS	Peppermint oil	Global improvement	4 weeks	Improvement in global symptoms in IBS		28/29	18/28	10/29	RR 1.86 (1.05,3.31)	NR	*Favours intervention*	High
	Dew 1984	IBS	Peppermint oil	Global improvement	2 weeks	Improvement in global symptoms in IBS		29	24/29	5/29	RR 4.8 (2.13, 10.04)	NR	*Favours intervention*	Some concerns

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Alammar 2019 cont'd	Lech 1988	IBS	Peppermint oil	Global improvement	4 weeks	Improvement in global symptoms in IBS		23/24	13/23	6/24	RR 2.26 (1.04, 4.93)	NR	*Favours intervention*	High
		IBS	Peppermint oil	Abdominal pain		Proportion of patients with improvement			12/23	6/24	RR 2.04 (0.94, 4.63)	NR	*Favours intervention*	High
	Liu 1997	IBS	Peppermint oil	Abdominal pain	4 weeks	Proportion of patients with improvement		55/55	41/55	21/55	RR 1.95 (1.35, 2.83)	NR	*Favours intervention*	Some concerns
	Merat 2010	IBS	Peppermint oil	Abdominal pain	8 weeks	Proportion of patients with improvement		45/45	19/45	16/45	RR 1.19 (0.71, 2.00)	NR	*Favours intervention*	Some concerns
	Weiss 1988	IBS	Peppermint oil	Global improvement	3 weeks	Improvement in global symptoms in IBS		30/30	17/30	4/30	RR 4.25 (1.62, 11.15)	NR	*Favours intervention*	Some concerns
	Schneider 1990	IBS	Peppermint oil	Abdominal pain	6 weeks	Proportion of patients with improvement		30/30	19/30	11/30	RR 1.73 (1.00, 2.97)	NR	*Favours intervention*	High
	Rees 1979	IBS	Peppermint oil	Global improvement	3 weeks	Improvement in global symptoms in IBS		18/18	13/18	5/18	RR 2.6 (1.17, 5.78)	NR	*Favours intervention*	High
	Footnotes:													
Hong 2018**	Davis 2006	IBS	Aloe vera	Global improvement	4 weeks	Improvement of ≥50 in IBS-SSS (mean)		26/23*	39.12 (77.45)	13.74 (85.03)	SMD 0.31 (-0.26, 0.87)	NR	*Favours intervention*	Moderate
	Hutchings 2011	IBS	Aloe vera	Global improvement	20 weeks	Improvement in global symptoms in IBS		12/13*	3.5 (2.26)	2.49 (1.7)	SMD 0.49 (-0.31, 1.29)	NR	*Favours intervention*	High
	Storsrud 2015	IBS	Aloe vera	Global improvement	4 weeks	Improvement of ≥50 in IBS-SSS (mean)		32/31*	58 (76.35)	23 (73.1)	SMD 0.46 (-0.04, 0.96)	NR	*Favours intervention*	Some concerns
	Footnotes:	*data presented are per protocol numbers; ** Only RCT data not included elsewhere are extracted here												

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Ng 2018	Brinkhaus 2005	IBS	Curcumin	Global improvement	18 weeks	Proportion of patients with improvement		106	NR	NR	NR	NR	*No difference*	Moderate
	Portincasa 2016	IBS	Curcumin	Global improvement	4 weeks	Changes in IBS-SSS		121	NR	NR	NR	<0.001	*Favours intervention*	Some concerns
				HRQoL		IBS quality of life			NR	NR	NR	< 0.001	*Favours intervention*	Some concerns
	Footnotes:													
Anheyer 2017	Kline 2001	IBS	Peppermint oil	Global improvement	2 weeks	GI symptom rating scale		42	NR	NR	NR	NR	*No difference*	Some concerns
	Shulman 2016	IBS	Pysllium fibre	Abdominal pain	6 weeks	Number pain episodes		103	NR	NR	NR	NR	*Favours intervention*	Some concerns
	Footnotes:													
WHM vs inactive control														
No RCTs reporting this comparison were identified by the eligible reviews														
WHM vs active control														
Black 2020*	Nigam 1984	IBS	Ispaghula husk OR Hyoscine OR Amitriptyline	Global improvement	Individual study data are not reported.									
	Ritchien 1979	IBS	Lorazepam with hyoscine OR ispaghula husk	Global improvement	Individual study data are not reported.									
	Footnotes:	*The authors conduct and report network meta-analysis results. Individual study data are not reported.												

Abbreviations: C, Comparator; Confidence interval; I, intervention; IBS, irritable bowel syndrome; IBS-SSS, IBS symptom severity scale; NR, not reported; RoB, risk of bias; SD, standard deviation; QoL, quality of life

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
WHM vs placebo														
Sadeghi 2020	Moeini 2016	GORD	Hawthorn	improvement of GORD symptoms	4 weeks	NR	NR	39/41	NR	NR	NR	NR	*Favours intervention*	Moderate
				Regurgitation		NR	NR		NR	NR	NR	NR	*Favours intervention*	Moderate
	Footnotes:													
WHM vs inactive control														
No RCTs reporting this comparison were identified by the eligible reviews														
WHM vs active control														
No RCTs reporting this comparison were identified by the eligible reviews														

Abbreviations: C, Comparator; Confidence interval; I, intervention; NR, not reported; RoB, risk of bias; SD, standard deviation

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
WHM vs placebo														
Negi 2021*	Jenabi 2013	See details reported by Chen 2016												
	Rahnama 2012	See details reported by Chen 2016												
	Abadi 2020	Dysmenorrhoea	Ginger 250 mg QID x 3 days	Pain duration	NR	NR	NR	50/50	1.61 (0.64)	2.12 (0.81)	NR	0.052	*Favours intervention*	Some concerns
	Kashefi 2014	Dysmenorrhoea	Ginger 250 mg TID x 4 days	Pain severity	NR	VAS	Higher means more pain	47/45	3.08 (1.52)	6.95 (1.67)	MD -3.87 (-4.52, -3.22)	<0.001	*Favours intervention*	High
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Xu 2020*	Jenabi 2013	See details reported by Chen 2016												
	Rahnama 2012	See details reported by Chen 2016												
	Kashefi 2014	See details reported by Chen 2016												
	Jaafarpour 2015	Dysmenorrhoea	Cinnamon 420 mg TID x 3 days	Pain intensity	1 cycle	VAS	Higher means more pain	38/38	NR	NR	WMD 2.10 (1.70, 2.50)	NR	*Favours intervention*	Low
	Jahangirifar 2018	Dysmenorrhoea	Cinnamon 1000mg TID x 3 days	Pain intensity	2 cycles	VAS	Higher means more pain	30/28	NR	NR	WMD 1.60 (1.45, 1.75)	NR	*Favours intervention*	Low
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Mollazadeh 2019	Shahhosseini 2005	Heavy menstrual bleeding	Vitex drops	Menstrual blood loss	3 months	Higham score	Higher means greater blood loss	30/30	NR	NR	NR	NR	*No difference*	Low
	Shobheiri 2014*	Heavy menstrual bleeding	Vitex drops (group 2)	Menstrual blood loss	2 cycles	Higham score	Higher means greater blood loss	30/30	25.6 (11.4)	24.6 (13.5)	1.00 (-5.74, 5.75)	NR	*No difference*	High
	Footnotes:	Shobheiri 2014 also tested vitex vs mefenamic acid (group 3)												

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Pellow 2018	Jenabi 2013	See details reported by Chen 2016												
	Rahnama 2012	See details reported by Chen 2016												
	Kashefi 2014	See details reported by Chen 2016												
	Younesy 2014	Dysmenorrhoea	Fenugreek 900 mg TID	Pain intensity	2 cycles	VAS	Higher means more pain	51/50	NR	NR	NR	<0.001	*Favours intervention*	Unclear
	Heshmati 2016	Dysmenorrhoea	Peppermint 990 mg	Pain severity	3 cycles	SF-MPQ	Higher means more pain	46/44	NR	NR	NR	0.008	*Favours intervention*	Unclear
	Mirabi 2011	Dysmenorrhoea	Valerian	Pain intensity	2 cycles	VAS	Higher means more pain	51/49	NR	NR	NR	<0.001	*Favours intervention*	Unclear
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Chen 2016	Jenabi 2013	Dysmenorrhoea	Ginger 500mg TID	Pain severity	1 menstrual cycle	VAS	Higher means more pain	35/34	4.81 (1.7)	7.11 (1.12)	MD -2.30 (-2.98, -1.62)	0.001	*Favours intervention*	Some concerns
	Rahnama 2012*	Dysmenorrhoea	Ginger capsule 50 mg TID	Pain severity	2 menstrual cycles	VAS	Higher means more pain	59/46	4.6 (2.6)	6 (2.7)	MD -1.4 (-2.4, -0.4)	0.017	*Favours intervention*	Some concerns
	Kashefi 2014	Dysmenorrhoea	Ginger 50mg TID	Pain severity	2 menstrual cycles	VAS	Higher means more pain	47/45	3.1 (1.5)	7 (1.7)	MD -3.87 (-4.54, -3.2)	NR	*Favours intervention*	Some concerns
	Footnotes:	*it is unclear if the reported results are for patients following Protocol 1 or 2 within the study, or overall results												
	Jenabi 2013	See details reported by Chen 2016												
	Rahnama 2012	See details reported by Chen 2016												
	Kashefi 2014	See details reported by Chen 2016												

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Pattanittum 2016*	Akbari 2012	Dysmenorrhoea	Fenugreek	Pain intensity	2 menstrual cycles	VAS	Higher means more pain	51/50	3.3 (1.3)	6 (1.9)	MD -2.71 (-3.33, -2.09)	NR	*Favours intervention*	Low
	Akhavan Amjadi 2009	Dysmenorrhoea	Cinnamon	Pain intensity	2 menstrual cycles	0-3 scale	not clear	Data not extracted as details on the measure used was not clear.						
	Rahnama 2010	Dysmenorrhoea	Ginger 500 mg x 3 days	Pain intensity	2 menstrual cycles	VAS	Higher means more pain	37/41	NR	NR	NR	<0.01	*Favours intervention*	High
	Dolation 2010	Dysmenorrhoea	Valerian root	Pain intensity	2 menstrual cycles	VAS	Higher means more pain	51/49	2 (1.4)	4.4 (1.8)	MD -2.42 (-3.05, -1.79)	NR	*Favours intervention*	Some concerns
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Daily 2015*	Jenabi 2013	See details reported by Chen 2016												
	Rahnama 2012	See details reported by Chen 2016												
	Kashefi 2014	See details reported by Chen 2016												
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
WHM vs control (no treatment, waitlist, standard care)														
Pattanittum 2016	Jenabi 2010	Dysmenorrhoea	Chamomile tea vs no intervention	Pain intensity	3 months	SF-MPQ	Higher means more pain	40/40	5.94 (6.01)	7.10 (10.39)	NR	<0.001	*Favours intervention*	High
	Modaress 2011	Dysmenorrhoea	German chamomile vs no intervention (adjunct to mefenamic acid)	Pain severity	2 menstrual cycles	VAS	Higher means more pain	80/80	0.4 (0.9)	4.2 (2.1)	MD -3.73 (-4.23, -3.23)	NR	*Favours intervention*	Some concerns
	Footnotes:													

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Daily 2015	Gupta 2013	Dysmenorrhoea	Ginger 500 mg vs no intervention (adjunct to specified exercise)	Pain intensity	2 cycles	NRS (10 point)	Higher means more pain	34/30	2.91 (2.45)	4.13 (2.12)	NR	0.039	*Favours intervention*	High
	Footnotes:													
WHM vs active control														
Negi 2021*	Pakniat 2019	Dysmenorrhoea	Ginger vs mefenamic acid	Pain severity	NR	VAS	Higher means more pain	50/50	3.12 (1.28)	6 (0.7)	-2.88 (-3.28, -2.48)	<0.001	*Favours intervention*	Some concerns
	Rad 2018	Dysmenorrhoea	Ginger vs Novafen	Pain severity	NR	VAS	Higher means more pain	78/90	3.10 (2.69)	2.97 (2.69)	NR	<0.05	*Not reported*	Some concerns
	Ozgoli 2009	See details reported by Chen 2016												
	Shirvani 2015	See details reported by Chen 2016												
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Xu 2020*	Pakniat 2019	See details reported by Negi 2021												
	Footnotes:	*Only RCT data not included elsewhere are extracted here												
Pattanittum 2016*	Jenabi 2012	Dysmenorrhoea	Valerian vs mefenamic acid	Pain severity	2 menstrual cycles	0 to 10 pain scale	Higher means more pain	49/50	3.7 (1.3)	3.1 (1.7)	0.62 (0.03, 1.21)	NR	*Favours comparator*	High
	Kashefi 2014	Dysmenorrhoea	Ginger vs zinc sulphate	Pain severity	2 menstrual cycles	VAS	Higher means more pain	45/53	3.1 (1.5)	3.1 (1.2)	-0.04 (-0.59, 0.51)	NR	*No difference*	Some concerns
	Footnotes:	*Only RCT data not included elsewhere are extracted here												

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Chen 2016	Kashefi 2014	Dysmenorrhoea	Ginger 250mg TID vs Zinc	Pain severity	2 cycles	VAS	Higher means more pain	NR	NR	NR	NR	NR	*Not reported*	Low
	Ozgoli 2009	Dysmenorrhoea	Ginger 250 mg vs Ibuprofen	Pain severity	Unclear	VMS	Unclear	NR	NR	NR	NR	NR	*Not reported*	High
			Ginger 250 mg vs mefenamic acid	Pain severity	Unclear	VMS	Unclear	50/50	NR	NR	SMD -0.18 (-0.040, 0.04)	NR	*No difference*	High
	Shirvani 2015*	Dysmenorrhoea	Ginger 250 mg vs mefenamic acid	Worst pain severity	1 cycle	VAS	Higher means more pain	61/61	NR	NR	SMD 0.24 (-0.12, 0.59)	NR	*No difference*	High
				Pain duration	1 cycle	Days in pain	Days in pain	NR	NR	NR	NR	NR	*Not reported*	High
	Halder 2012	Dysmenorrhoea	Ginger 1000 mg BID vs PMR	Dysmenorrhoea severity	measured every 24 hours	5-point scale	Higher is worse pain	NR	NR	NR	NR	NR	*Not reported*	High
	Footnotes:	*there was differential use of extra analgesics between the ginger and NSAID groups in this study												
Daily 2015*	Shirvani 2015	See details reported by Chen 2016												
	Kashefi 2014	See details reported by Chen 2016												
	Ozgoli 2009	See details reported by Chen 2016												
	Halder 2012	See details reported by Chen 2016												
	Footnotes:	*Only RCT data not included elsewhere are extracted here												

Abbreviations: BID, twice daily; C, Comparator; Confidence interval; I, intervention; NR, not reported; PMR, progressive muscle relaxation' QID, four times daily; RoB, risk of bias; SD, standard deviation; TID, three times daily; WMD, weigthed mean difference;

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
WHM vs placebo														
Ghaderi 2020	Agha-Hosseini	PMS	Saffron (30mg/day)	Depressive symptoms	8 weeks	HDRS-D*	higher is worse	47 (24/23)	NR	NR	WMD 6.23 (5.21, 7.25)	NR	*Favours intervention*	Not reported
	Footnotes:	*Not clear if the primary study measure other PMS-related outcomes.												
Shinjyo 2020	Behboodi Moghadam 2016	PMS	Valerian	Anxiety	3 menstrual cycles	Self-rated emotional state (anxiety)	not specified*	100 (NR)	NR	NR	Hedges' g 1.9 (1.44, 2.39)	NR	*Favours intervention*	High
	Footnotes:	*The primary study measured PMS (emotional, behavioural and physical) symptoms, but only anxiety reported by the SR authors.												
Csupor 2019	He 2009	PMS	Chaste tree berry (40 mg)	Overal PMS symptoms	3 menstrual cycles	Response rate	≥ 60% increase in total PMSD score	208 (104/104)	83/104 (80%)	52/104 (50)	RR 1.22 (1.29, 1.98)	NR	*Favours intervention*	Some concerns
	Schellenberg 2001	PMS	Chaster tree berry 20 mg (ZE440)	Overal PMS symptoms	NR	Response rate	≥ 50% decrease in total TSS score	170 (86/84)	45/86 (52%)	20/84 (24%)	RR 2.20 (1.43, 3.39)	NR	*Favours intervention*	Some concerns
	Schellenberg 2012 *	PMS	Chaste tree berry (8mg)	Overal PMS symptoms	NR	Response rate	≥ 50% decrease in total TSS score	71 (36/35)	5/36 (14%)	4/35 (11%)	RR 1.22 (0.36, 4.16)	NR	*Favours intervention*	Some concerns
			Chaste tree berry (20mg)	Overal PMS symptoms	NR	Response rate	≥ 50% decrease in total TSS score	70 (35/35)	28/35 (81%)	4/35 (11%)	RR 7.00 (2.74, 17.87)	NR	*Favours intervention*	Some concerns
			Chaste tree berry (30mg)	Overal PMS symptoms	NR	Response rate	≥ 50% decrease in total TSS score	71 (36/35)	22/36 (61%)	4/35 (11%)	RR 5.35 (2.05, 13.94)	NR	*Favours intervention*	Some concerns
	Csupor 2019		**meta-analysis results***		**Response rate**		**3 studies**	**Random effects**	**RR 2.57; 95% CI 1.52, 4.35; p=0.00004; I2=75%**				*Favours intervention*	
	Footnotes:	*SR authors report the Schellenberg 2012 data separate - but placebo group is the same for each arm.												

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Khalesi 2019	Najafi 2018	PMS	Chamomile	Overal PMS symptoms	NR	NR	NR	118 (NR)	SR authors did not report any data				*Favours intervention*	Not reported
	Footnotes:													
Verkaik 2017*	He 2009	PMS	Chaste tree berry 40 mg (BNO1095)	Overal PMS symptoms	3 menstrual cycles	PMSD	higher is worse	217 (108/109)	NR	NR	-0.81 (-1.10, -.52)*	NR	*Favours intervention*	Some concerns
	Kaplanoglu 2015	PMS	Chaste tree berry	Overal PMS symptoms	3 menstrual cycles	PMS-VAS	higher is worse	120 (40/40/40)	NR	NR	-0.75 (-1.2, -3.0)*	NR	*Favours intervention*	High
				Anxiety		VAS	higher is worse		NR	NR	-1.71 (-2.22, -1.20)*	NR	*Favours intervention*	High
				Depressive symptoms		VAS	higher is worse		NR	NR	-1.12 (-1.58, -0.65)*	NR	*Favours intervention*	High
	Mousavi 2015	PMS	Chaste tree berry	Overal PMS symptoms	3 menstrual cycles	PMS-VAS	higher is worse	72 (36/36)	NR	NR	-2.38 (-3.03, -1.73)*	NR	*Favours intervention*	High
				Depressive symptoms		VAS	higher is worse		NR	NR	-2.30 (-2.95, -1.66)	NR	*Favours intervention*	High
	Pakgohar 2009	PMS	Chaste tree berry	Overal PMS symptoms	2 menstrual cycles	Daily symptom rating	higher is worse	116 (49/50)	NR	NR	-1.21 (-1.63, -0.78)*	NR	*Favours intervention*	Low
				Depressive symptoms		BDI	higher is worse		NR	NR	-1.07 (-1.48, -0.65)	NR	*Favours intervention*	Low
	Risoleti 2011	PMS	Chaste tree berry	Overal PMS symptoms	3 menstrual cycles	PMSD	higher is worse	72 (27/20)	NR	NR	-0.89 (-1.49, -029)*	NR	*Favours intervention*	Low
	Schellenberg 2001	PMS	Chaste tree berry 20 mg (ZE440)	Overal PMS symptoms	3 menstrual cycles	PMS-VAS	higher is worse	(178/170)	NR	NR	-0.28 (-0.58, 0.02)*	NR	*No difference*	High

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Verkaik 2017* cont'd	Schellenberg 2012**	PMS	Chaste tree berry (8mg)	Overal PMS symptoms	3 menstrual cycles	PMS-VAS	higher is worse	162 (36/35/36/35)	NR	NR	-0.20 (-0.66, 0.26)*	NR	*No difference*	Some concerns
			Chaste tree berry (20mg)	Overal PMS symptoms			higher is worse		NR	NR	-3.36 (-4.08, -2.64)*	NR	*Favours intervention*	Some concerns
			Chaste tree berry (30mg)	Overal PMS symptoms			higher is worse		NR	NR	-2.66 (-3.30, -2.03)*	NR	*Favours intervention*	Some concerns
	Zamani 2012	PMS	Chaste tree berry	Overal PMS symptoms	6 menstrual cycles	PMS-VAS	higher is worse	134 (62/66)	NR	NR	-1.00 (-1.37, -0.64)*	NR	*Favours intervention*	Some concerns
				Anxiety (nervousness)		VAS (0-10)	higher is worse		NR	NR	-1.23 (-1.61, -0.85)*	NR	*Favours intervention*	Some concerns
				Depressive symptoms		VAS (0-10)	higher is worse		NR	NR	-0.77 (-1.13, -0.41)	NR	*Favours intervention*	Some concerns
	Turner 1993	PMS	Chaste tree berry	Depressive symptoms	3 menstrual cycles	MDQ negative affect	higher is worse	217 (105, 112)	NR	NR	-0.06 (-0.32, 0.21)	NR	*No difference*	High
	Verkaik 2017		**meta-analysis results****		**Overal PMS symptoms**		**8 studies**	**N=NR**	**SMD -1.31; 95% CI -1.82, -0.80; p=0.000; I2=92.6%**				*Favours intervention*	
					Depressive symptoms		**5 studies**	**N=NR**	**SMD -1.02; 95% CI -1.67, -0.38; p=0.000; I2=92.4%**				*Favours intervention*	
					Anxiety symptoms		**2 studies**	**N=NR**	**SMD -1.44; 95% CI -1.91, -0.97; p=0.137; I2=54.9%**				*Favours intervention*	
	Footnotes:	*Data presented as Hedge's g (95% CI) **SR authors report the Schellenberg 2012 data separate - but placebo group is the same for each arm.												
van Die 2013	Zamani 2012#	PMS	Chaste tree berry	Depression	6 menstrual cycles	VAS (0-10)	higher is worse	134 (62/66)	2.8 (1.8)	6.0 (1.5)	NR	<0.0001	*Favours intervention*	Low
				Anxiety (nervousness)		VAS (0-10)	higher is worse		3.8 (2.0)	5.7 (2.1)	NR	<0.0001	*Favours intervention*	Low

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
van Die 2013 cont'd	Ma 2010	PMS	Chaste tree berry 40 mg (BNO1095)	Overal PMS symptoms	3 menstrual cycles	PMSD	NR	67 (33/34)	4.28 (5.76)	11.79 (11.78)	NR	< 0.05	*Favours intervention*	Low
				Overal PMS symptoms		PMTS	NR		8.71 (8.62)	14.44 (10.64)	NR	< 0.05	*Favours intervention*	Low
				Overal PMS symptoms		Response rate	response rate (% improved)		84.85%	55.89%	NR	NR	*Favours intervention*	Some concerns
	He 2009	PMS	Chaste tree berry 40 mg (BNO1095)	Overal PMS symptoms	3 menstrual cycles	PMSD	NR	217 (108/109)	6.41 (7.94)	12.64 (10.35)	NR	<0.0001	*Favours intervention*	Low
				Overal PMS symptoms		PMTS	NR		9.92 (9.01)	14.59 (10.69)	NR	< 0.05	*Favours intervention*	Low
				Overal PMS symptoms		Response rate	response rate (% improved)		79.80%	50.00%	NR	NR	*Favours intervention*	Some concerns
	Pakgohar 2009	PMS	Chaste tree berry 4.3-4.8 mg	Overal PMS symptoms	2 menstrual cycles	Daily symptom rating	higher is worse	116 (58/58)	13.28 (10.82)	24.57 (12.42)	NR	NR	*Favours intervention*	Some concerns
				Overal PMS symptoms		Daily symptom rating	response rate (% improved)		60.73%	20.79%	NR	NR	*Favours intervention*	Some concerns
				Psychological symptoms		DSR	higher is worse		12.01 (12.22)	21.73 (14.44)	NR	NR	*Favours intervention*	Some concerns
				Psychological symptoms		DSR	response rate (% improved)		65.62%	28.19%	NR	NR	*Favours intervention*	Some concerns
				Physical symptoms		DSR	higher is worse		12.63 (10.62)	24.33 (12.02)	NR	NR	*Favours intervention*	Some concerns
				Physical symptoms		DSR	response rate (% improved)		57.98%	16.22%	NR	NR	*Favours intervention*	Some concerns

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
van Die 2013 cont'd	Schellenberg 2001	PMS	Chaste tree berry 20 mg (ZE440)	Overal PMS symptoms **	3 menstrual cycles	VAS (0-10)	higher is worse	178 (91/87)	134.5*	177.9*	NR	NR	*Favours intervention*	Low
				Overal PMS symptoms **		CGI severity	higher is worse		3.7*	4.02*	NR	NR	*Favours intervention*	Low
				Overal PMS symptoms **		Response rate	≥ 50% decrease in total TSS score		52%	24%	NR	NR	*Favours intervention*	Some concerns
	Turner 1993	PMS	Chaste tree berry 600mg tds	Overal PMS symptoms	3 menstrual cycles	Moos MDQ	higher is worse	NR	NR	NR	NR	NR	*Not reported*	High
				Overal PMS symptoms			response rate (% improved)		25/62	16/74	NR	NR	*Favours comparator*	High
	Footnotes:	#study also assessed headaches, restlessness, breast pain, bloating/tympanis (not extracted here [not critical or imporant outcomes]) *SD not reported ** inclusive of irritability, mood, anger, headache, bloating, breast fullnesss												
WHM vs control (no treatment, waitlist, standard care)														
No RCTs reporting this comparison were identified by the eligible reviews														
WHM vs 'other'														
nalesi 2019	Sharifi 2014	PMS	Chamomile vs Mefenamic Acid (250 mg)	Emotional symptoms	NR	NR	NR	90 (NR)	SR authors did not report any data				*Favours intervention*	Not reported
	Karimian 2013	PMS	Chamomile vs Mefenamic Acid (250 mg)	Physical symptoms	NR	NR	NR	90 (NR)	SR authors did not report any data				*Favours intervention*	Not reported

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Kh	Modaress 2011	PMS	Chamomile vs Mefenamic Acid (250 mg)	Physical symptoms	NR	NR	NR	80 (NR)	SR authors did not report any data				*Favours intervention*	Not reported
	Footnotes:													
Verkaik 2017*	Atmaca 2003	PMS	Chaste tree vs fluoxetine	Overal PMS symptoms	3 menstrual cycles	Daily symptom	NR	19/19	NR	NR	0.05 (-0.57, 0.68)*	NR	*No difference*	Low
				Depressive symptoms		HAM-D	NR		NR	NR	0.29 (-0.34, 0.92)	NR	*No difference*	Low
	Ciotta 2011	PMDD	Chaste tree vs fluoxetine	Depressive symptoms	2 menstrual cycles	4 items of HAM-D	higher is worse	31/26	NR	NR	NR	NR	*Favours comparator*	High
	Di Pierro 2009	PMS	Chaste tree berry 40 mg vs magnesium 300 mg /day	Overal PMS symptoms	3 menstrual cycles	VAS on 8 symptoms	NR	82 (42/40)	NR	NR	-3.68 (-4.39, -2.97)*	NR	*Favours intervention*	High
	Kaplanoglu 2015	PMS	Chaste tree vs oral contraceptive	Overal PMS symptoms	3 menstrual cycles	VAS on 15 symptoms	higher is worse	120 (40/40/40)	NR	NR	-0.21 (-0.65, 0.22)*	NR	*Favours intervention*	Some concerns
				Depressive symptoms		VAS	NR		NR	NR	-0.75 (-1.20, -0.30)*	NR	*Favours intervention*	Some concerns
				Anxiety		VAS	higher is worse		NR	NR	0.35 (-0.09, 0.78)	NR	*Favours intervention*	Some concerns
	Lauritzen 1997	PMS	Chaste tree vs pyridoxine (Vit B6)	Overal PMS symptoms	3 menstrual cycles	PMTS-Response rate	≥10 point improvement	127 (61/66)	NR	NR	0.50 (-0.85, -0.15)	NR	*Favours intervention*	Some concerns
	Onaran 2003	PMS	Chaste tree vs oral contraceptive	Overal PMS symptoms	3 menstrual cycles	COPE	NR	124 (61/63)	NR	NR	0.01 (-0.36, 0.34)	NR	*No difference*	High
				Anxiety		HADS	higher is worse		NR	NR	0.00 (-0.58, 0.49)	NR	*No difference*	High
				Depressive symptoms		HADS	NR		NR	NR	-0.18 (-0.53, 0.17)	NR	*Favours intervention*	High

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
	Risoleti 2011	PMS	Chaste tree vs oral contraceptive	Overal PMS symptoms	3 menstrual cycles	PMSD	NR	72 (27/25)	NR	NR	-0.04 (-0.58, 0.49)	NR	*Favours intervention*	High
	Salehi 2013	PMS	Chaste tree vs Vitamin E	Overal PMS symptoms	2 cycles	PMSD	NR	225 (70/70)	NR	NR	-0.83 (-1.17, -0.49)	NR	*Favours intervention*	Some concerns
	Scaldarella 2008	PMS	Chaste tree vs pyridoxine (Vit B6)	Overal PMS symptoms	3 cycles	PMTS	NR	60 (30/30)	NR	NR	NR	NR	*Favours intervention*	High
	Footnotes:	*Data presented as Hedge's g (95% CI) (adjusted for small study bias)												
van Die 2013	Atmaca 2003	PMDD	Chaste tree berry ? mg vs fluoxetine (20-40 mg/day)	Overal PMS symptoms	3 menstrual cycles	Daily symptom rating	higher is worse	41 (20/21)	82.8 (49.5)	85.6 (55.3)	0.05 (-0.57, 0.68)*	NR	*No difference*	High
				Depressive symptoms		HAM-D	higher is worse		7.6 (4.3)	7.1 (3.8)	0.29 (-0.34, 0.92)	NR	*No difference*	High
				Efficacy		CGI-Severity	higher is worse		1.2 (0.7)	1.5 (0.6)	NR	NR	*No difference*	High
				Efficacy		CGI-Severity	% improved		57.90%	68.40%	NR	NR	*Not reported*	High
	Ciotta 2011	PMDD	Chaste tree berry 20 mg vs fluoxetine (20-40 mg/day)	Depressive symptoms	2 menstrual cycles	HAM-D	higher is worse	57 (31/26)	69	36	NR	< 0.02	*Favours comparator*	Low
				Work interest		HAM-D	higher is worse		58	26	NR	<0.05	*Favours comparator*	Low
				Psychic anxiety		HAM-D	higher is worse		80	22	NR	<0.01	*Favours comparator*	Low
				General somatic		HAM-D	higher is worse		37	14	NR	< 0.05	*Favours comparator*	Low
	Lauritzen 1997	PMS	Chaste tree berry 3.5-4.2 mg vs pyridoxine B6	Overal PMS symptoms	3 menstrual cycles	PMTS	higher is worse	175 (61/66)	5.1 (6.6)	5.1 (6.6)	NR	NR	*Favours intervention*	Low
				Overal PMS symptoms		CGI	% 'excellent'		24.50%	12.10%	NR	NR	*Not reported*	Some concerns

STUDY RESULTS (as reported by review authors)

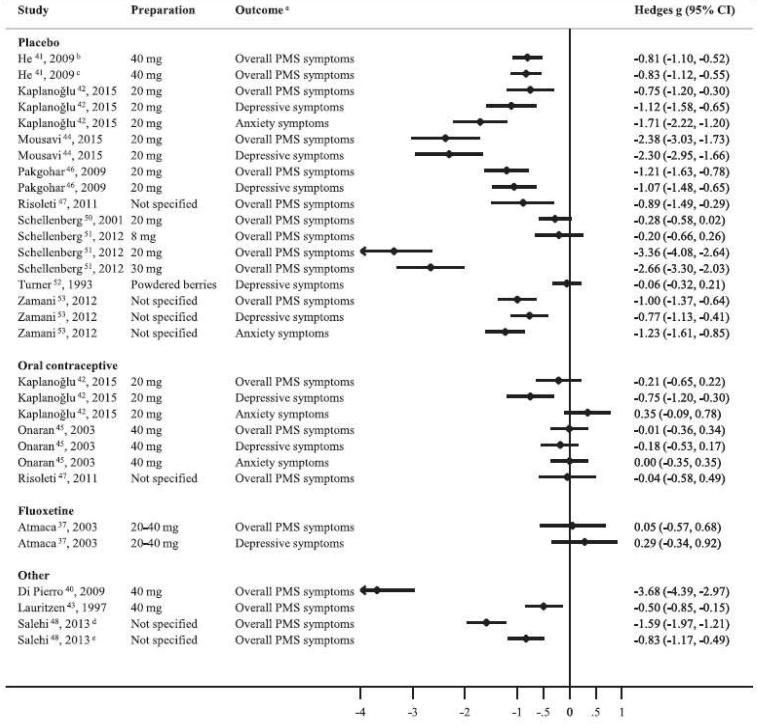
Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
			pyridoxine B6 (100 mg, bid)	Overal PMS symptoms	cycles	Response rate	% free from complaints		36.10%	21.30%	NR	NR	*Not reported*	Some concerns

STUDY RESULTS (as reported by review authors)														
Review ID	**Study ID**	**Condition**	**Intervention**	**Outcome**	**Timing**	**Outcome measure**	**measure details**	**# participants (I/C)**	**(intervention) n/N (%) or mean (SD)**	**(comparator) n/N (%) or mean (SD)**	**Point estimate (95% CI)**	***p* -value**	***direction of effect***	**RoB**
van Die 2013 cont'd	Di Pierro 2009	PMS	Chaste tree berry 40 mg vs magnesium 300 mg /day	Back pain	6 menstrual cycles	VAS (0-10)	higher is worse	82 (42/40)	1.4 (0.6)	5.5 (1.5)	NR	< 0.001	*Favours intervention*	High
				Menstrual pain		VAS (0-10)	higher is worse		2.0 (0.4)	7.4 (2.6)	NR	< 0.001	*Favours intervention*	High
				Breast fullness		VAS (0-10)	higher is worse		0.4 (0.6)	5.4 (1.8)	NR	< 0.001	*Favours intervention*	High
				Headache		VAS (0-10)	higher is worse		1.0 (0.8)	6.4 (1.6)	NR	< 0.001	*Favours intervention*	High
				Asthenia		VAS (0-10)	higher is worse		1.2 (0.9)	6.4 (1.2)	NR	< 0.001	*Favours intervention*	High
				irritability		VAS (0-10)	higher is worse		0.4 (0.4)	5.0 (1.1)	NR	< 0.001	*Favours intervention*	High
				sleep		VAS (0-10)	higher is worse		1.1 (0.5)	5.4 (0.7)	NR	< 0.001	*Favours intervention*	High
				appetite modulation		VAS (0-10)	higher is worse		0.2 (0.2)	1.4 (0.6)	NR	NS	*No difference*	High
	Footnotes:													

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Verkaik 2017	Results from Verkaik 2017													

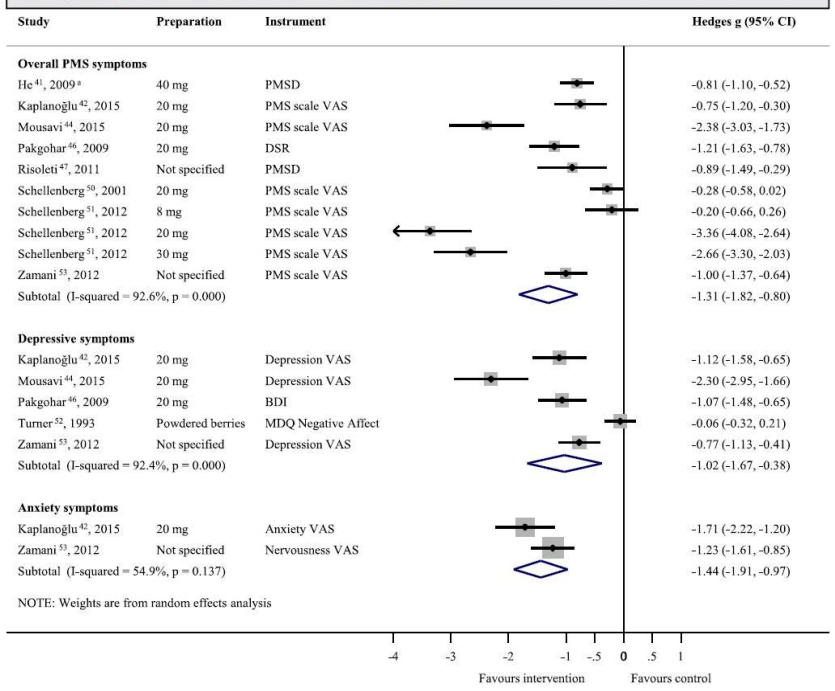
FIGURE 3
Forest plot of included studies stratified by comparator



Treatment effects of *Vitex agnus castus* (N=14; 31 effect sizes) are expressed in Hedges g and 95% confidence intervals (CI). The solid vertical line corresponds to no effect of treatment. [a]Overall premenstrual syndrome (PMS) symptoms are assessed using the following instruments: PMS diary, Premenstrual Tension Self-Rating Scale, Daily Symptom Report, Calendar of Premenstrual Experiences, and PMS visual analog scale; Depressive symptoms are assessed using Beck Depression Inventory, Hamilton Depression Rating Scale, Mood Disorders Questionnaire, or visual analog scale; Anxiety symptoms are assessed using anxiety subscale of the Hamilton Depression Rating Scale, or visual analog scale. [b]Assessed using PMS diary. [c]Premenstrual Tension Self-Rating Scale. [d]*Hypericum Perforatum* used as control. [e]Vitamin E used as control.

Verkaik. Treatment of premenstrual syndrome with preparations of Vitex agnus castus. Am J Obstet Gynecol 2017.

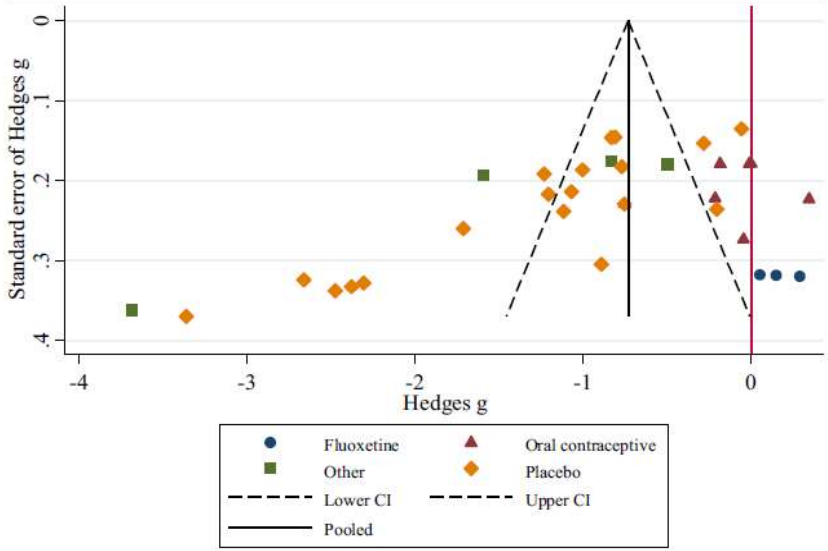
FIGURE 4
Forest plot of placebo-controlled studies stratified by outcome



STUDY RESULTS (as reported by review authors)

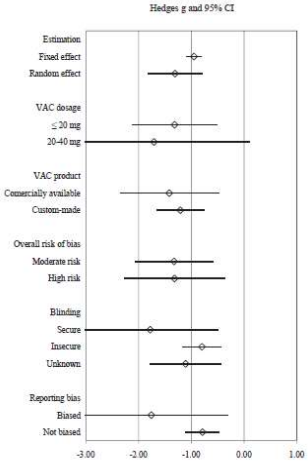
Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Verkaik 2017	Results from Verkaik 2017													

FIGURE 5
Funnel plot of included studies stratified by comparator



Verkaik. Treatment of premenstrual syndrome with preparations of Vitex agnus castus. Am J Obstet Gynecol 2017.

FIGURE 6
Vitex agnus castus treatment effect for different subgroups of studies



Treatment effect of *Vitex agnus castus* (*VAC*) of placebo-controlled studies with overall PMS symptoms as outcome (N=8; 10 effect sizes) using fixed and random effect estimation, and for different subgroups of studies (N=9). Pooled effect sizes for subgroups of studies are estimated using random effects estimation.

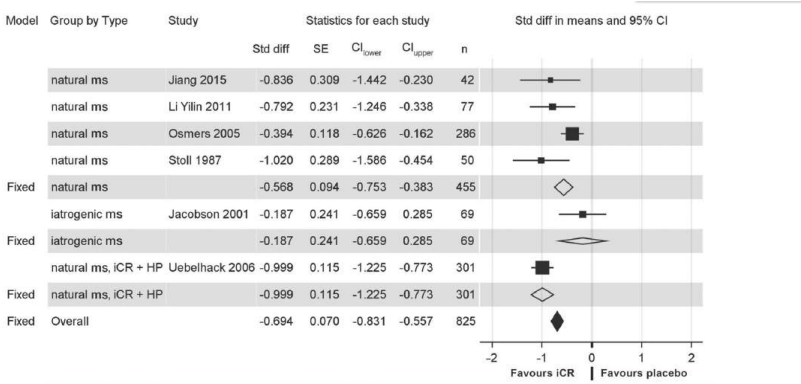
CI, confidence interval.

Verkaik. Treatment of premenstrual syndrome with preparations of Vitex agnus castus. Am J Obstet Gynecol 2017.

Abbreviations: C, Comparator; COPE, Calendar of Premenstrual Experiences; CGI-SI, Clinical Global Impression Scale; DSR, daily symptom report; I, intervention; IBS, irritable bowel syndrome; HDRS-D, Hamilton Depression Rating Scale; MDQ, Menstrual Distress Questionnaire; NR, not reported; PMS; premenstrual symdrome; PMSD, premenstrual symptom diary; TSS, total symptom score;

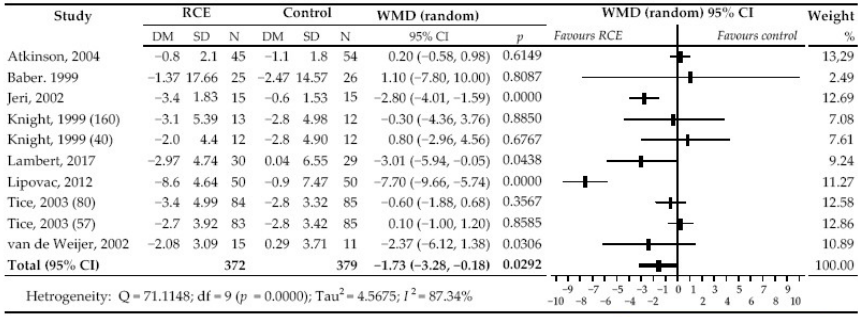
STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
WHM vs placebo														
Castelo-Branco 2021*	Jiang 2015	Menopausal symptoms	Black cohosh extract	Quality of life	6 months	MenQoL-physical	Higher is worse	24/24	NR	NR	SMD -0.836 (-1.442, -0.230)	NR	*Favours intervention*	Some concerns
	Li 2011	Menopausal symptoms	Black cohosh extract	Symptom severity	3 months	KMI	Higher is worse	45/32	NR	NR	SMD -0.792 (-1.246, -0.338)	NR	*Favours intervention*	High
				Vasomotor symptoms	3 months	Hot flashes, daily frequency	Higher is worse	45/32	NR	NR	NR	NR	*Favours intervention*	High
	Osmers 2005	Menopausal symptoms	Black cohosh extract	Symptom severity	3 months	MRS total	Higher is worse	153/151	NR	NR	SMD -0.394 (-0.626, -0.162)	NR	*Favours intervention*	Low
				Vasomotor symptoms	3 months	Hot flashes, daily frequency	Higher is worse	153/151	NR	NR	NR	NR	*Favours intervention*	Low
	Stoll 1987	Menopausal symptoms	Black cohosh extract	Symptom severity	3 months	KMI	Higher is worse	30/20	NR	NR	SMD -1.020 (-1.586, -0.454)	NR	*Favours intervention*	Low
				Anxiety	3 months	HAM-A	Higher is worse	30/20	NR	NR	NR	NR	*Favours intervention*	Low
	Jacobson 2001	Menopausal symptoms	Black cohosh extract	Vasomotor symptoms	2 months	Hot flashes, daily frequency	Higher is worse	42/43	NR	NR	SMD -0.187 (-0.659, 0.285)	NR	*No difference*	Some concerns
	Uelelhack 2006	Menopausal symptoms	Black cohosh extract + St John's Wort	Symptom severity	4 months	MRS total	Higher is worse	151/150	NR	NR	SMD -0.999 (-1.225, -0.773)	NR	*Favours intervention*	Low
				Depression	4 months	HAM-D	Higher is worse	151/150	NR	NR	NR	NR	*Favours intervention*	Low

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
	Footnotes:	*fixed effects model	 Figure 2. Forest plot of isopropanolic *Cimicifuga racemosa* extract (iCR) versus placebo in neurovegetative and psychological menopausal symptoms. CI, confidence interval; Fixed, summary of the respective group of studies under the fixed-effect size model; HP, *Hypericum perforatum* (St. John's wort); ms, menopausal symptoms; n, number of patients; Std diff, standardized mean difference; SE, standard error of the standardized difference.											
Firoozeei 2021*	Kamalifard 2017	Menopausal symptoms	Oral lavender	Depression	8 weeks	BDI	Higher is worse	156 (NR)	NR	NR	SMD -1.12 (-153, -0.70)	NR	*Favours intervention*	Low
	Footnotes:	*random effects model. Included studies in other conditions. Only RCT in perimenopausal women included here.												
Kanadys 2021	Knight 1999	Menopausal symptoms	Red clover	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency	Higher is worse	37 (NR)	Data reported in forest plots below					High
				Symptom severity	12 weeks	GCS	Higher is worse	37 (NR)						High
	Baber 1999	Menopausal symptoms	Red clover	Vasomotor symptoms	90 days	Hot flashes, daily frequency	Higher is worse	51 (NR)						High
				Symptom severity	90 days	GCS	Higher is worse	51 (NR)						High
	Jeri 2002	Menopausal symptoms	Red clover	Vasomotor symptoms	16 weeks	Hot flashes, daily frequency	Higher is worse	30 (NR)						High
	van de Weijer 2002	Menopausal symptoms	Red clover	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency	Higher is worse	30 (NR)						High
				Symptom severity	12 weeks	GCS	Higher is worse	30 (NR)						High

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Kanadys 2021 cont'd	Tice 2003	Menopausal symptoms	Red clover	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency	Higher is worse	252 (NR)	Data reported in forest plots below					Low
	Atkinson 2004	Menopausal symptoms	Red clover	Vasomotor symptoms	12 months	Hot flashes, daily frequency	Higher is worse	205 (NR)						Low
				Symptom severity	12 months	GCS	Higher is worse	205 (NR)						Low
	Hidalgo 2005	Menopausal symptoms	Red clover	Vasomotor symptoms	90 days	Hot flashes, daily frequency	Higher is worse	60 (NR)						High
				Symptom severity	90 days	KMI	Higher is worse	60 (NR)						High
	del Giorno 2010	Menopausal symptoms	Red clover	Vasomotor symptoms	12 months	Hot flashes, daily frequency	Higher is worse	120 (NR)						Low
				Symptom severity	12 months	KMI	Higher is worse	120 (NR)						Low
	Lipovac 2012	Menopausal symptoms	Red clover	Vasomotor symptoms	12 months	Hot flashes, daily frequency	Higher is worse	113 (NR)						Some concerns
				Symptom severity	12 months	KMI	Higher is worse	113 (NR)						Some concerns
	Clifton-Bligh 2015	Menopausal symptoms	Red clover	Vasomotor symptoms	2 years	Hot flashes, daily frequency	Higher is worse	147 (NR)						Low
				Symptom severity	2 years	GCS	Higher is worse	147 (NR)						Low
	Shakeri 2015	Menopausal symptoms	Red clover	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency	Higher is worse	72 (NR)						Low
				Symptom severity	12 weeks	MRS total	Higher is worse	72 (NR)						Low
	Lambert 2017	Menopausal symptoms	Red clover	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency	Higher is worse	62 (NR)						Low
				Symptom severity	12 weeks	GCS	Higher is worse	62 (NR)						Low

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Kanadys 2021 cont'd														
	Footnotes:													

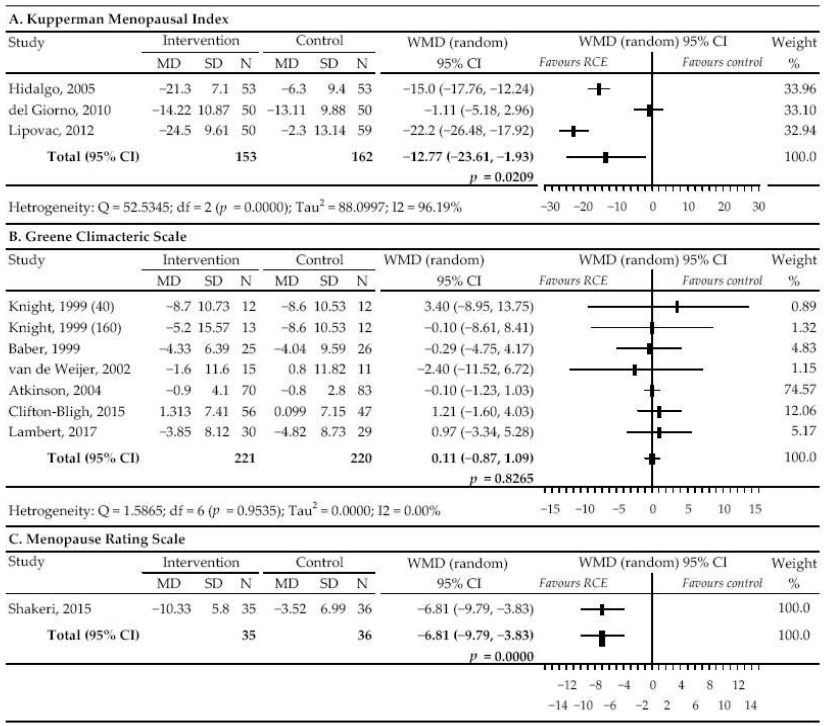


Study	RCE DM	RCE SD	RCE N	Control DM	Control SD	Control N	WMD (random) 95% CI	p	Weight %
Atkinson, 2004	−0.8	2.1	45	−1.1	1.8	54	0.20 (−0.58, 0.98)	0.6149	13.29
Baber, 1999	−1.37	17.66	25	−2.47	14.57	26	1.10 (−7.80, 10.00)	0.8087	2.49
Jeri, 2002	−3.4	1.83	15	−0.6	1.53	15	−2.80 (−4.01, −1.59)	0.0000	12.69
Knight, 1999 (160)	−3.1	5.39	13	−2.8	4.98	12	−0.30 (−4.36, 3.76)	0.8850	7.08
Knight, 1999 (40)	−2.0	4.4	12	−2.8	4.90	12	0.80 (−2.96, 4.56)	0.6767	7.61
Lambert, 2017	−2.97	4.74	30	0.04	6.55	29	−3.01 (−5.94, −0.05)	0.0438	9.24
Lipovac, 2012	−8.6	4.64	50	−0.9	7.47	50	−7.70 (−9.66, −5.74)	0.0000	11.27
Tice, 2003 (80)	−3.4	4.99	84	−2.8	3.32	85	−0.60 (−1.88, 0.68)	0.3567	12.58
Tice, 2003 (57)	−2.7	3.92	83	−2.8	3.42	85	0.10 (−1.00, 1.20)	0.8585	12.86
van de Weijer, 2002	−2.08	3.09	15	0.29	3.71	11	−2.37 (−6.12, 1.38)	0.0306	10.89
Total (95% CI)			**372**			**379**	**−1.73 (−3.28, −0.18)**	**0.0292**	100.00

Hetrogeneity: Q = 71.1148; df = 9 (p = 0.0000); Tau^2 = 4.5675; I^2 = 87.34%

WMD (random) 95% CI: *Favours RCE* / *Favours control*

Figure 4. Effects of isoflavones with red clover (*Trifolium pratense*) vs. placebo on the daily frequency of hot flushes in peri- and post-menopausal women. Number in brackets following author's name refers to dose of isoflavones in the study with more than one active group [33–38,41,44]. Abbreviations: RCIE, red clover isoflavone extract; WMD, weighted mean difference.



A. Kupperman Menopausal Index

Study	Intervention MD	Intervention SD	Intervention N	Control MD	Control SD	Control N	WMD (random) 95% CI	Weight %
Hidalgo, 2005	−21.3	7.1	53	−6.3	9.4	53	−15.0 (−17.76, −12.24)	33.96
del Giorno, 2010	−14.22	10.87	50	−13.11	9.88	50	−1.11 (−5.18, 2.96)	33.10
Lipovac, 2012	−24.5	9.61	50	−2.3	13.14	59	−22.2 (−26.48, −17.92)	32.94
Total (95% CI)			**153**			**162**	**−12.77 (−23.61, −1.93)**	100.0
							p = 0.0209	

Hetrogeneity: Q = 52.5345; df = 2 (p = 0.0000); Tau^2 = 88.0997; I2 = 96.19%

B. Greene Climacteric Scale

Study	Intervention MD	Intervention SD	Intervention N	Control MD	Control SD	Control N	WMD (random) 95% CI	Weight %
Knight, 1999 (40)	−8.7	10.73	12	−8.6	10.53	12	3.40 (−8.95, 13.75)	0.89
Knight, 1999 (160)	−5.2	15.57	13	−8.6	10.53	12	−0.10 (−8.61, 8.41)	1.32
Baber, 1999	−4.33	6.39	25	−4.04	9.59	26	−0.29 (−4.75, 4.17)	4.83
van de Weijer, 2002	−1.6	11.6	15	0.8	11.82	11	−2.40 (−11.52, 6.72)	1.15
Atkinson, 2004	−0.9	4.1	70	−0.8	2.8	83	−0.10 (−1.23, 1.03)	74.57
Clifton-Bligh, 2015	1.313	7.41	56	0.099	7.15	47	1.21 (−1.60, 4.03)	12.06
Lambert, 2017	−3.85	8.12	30	−4.82	8.73	29	0.97 (−3.34, 5.28)	5.17
Total (95% CI)			**221**			**220**	**0.11 (−0.87, 1.09)**	100.0
							p = 0.8265	

Hetrogeneity: Q = 1.5865; df = 6 (p = 0.9535); Tau^2 = 0.0000; I2 = 0.00%

C. Menopause Rating Scale

Study	Intervention MD	Intervention SD	Intervention N	Control MD	Control SD	Control N	WMD (random) 95% CI	Weight %
Shakeri, 2015	−10.33	5.8	35	−3.52	6.99	36	−6.81 (−9.79, −3.83)	100.0
Total (95% CI)			**35**			**36**	**−6.81 (−9.79, −3.83)**	100.0
							p = 0.0000	

WMD (random) 95% CI: *Favours RCE* / *Favours control*

Figure 5. Effects of red clover (*Trifolium pratense*) isoflavones vs. placebo on rating menopausal symptoms using the

Ghaderi 2020	Kashani 2018	Menopausal symptoms	Saffron	Depression	6 weeks	HAM-D	Higher is worse	28/28	NR	NR	MD 3.38 (2.05, 4.71)	NR	*Favours comparator*	Not reported
	Footnotes:													

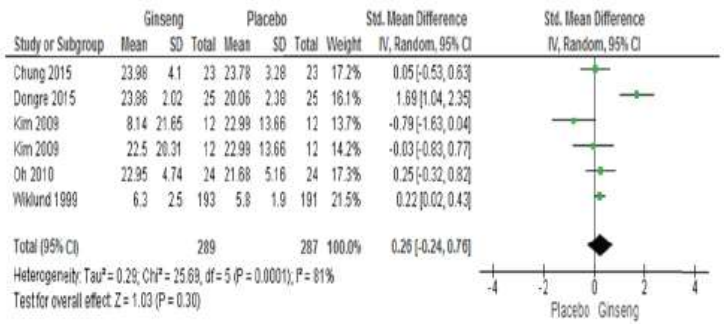
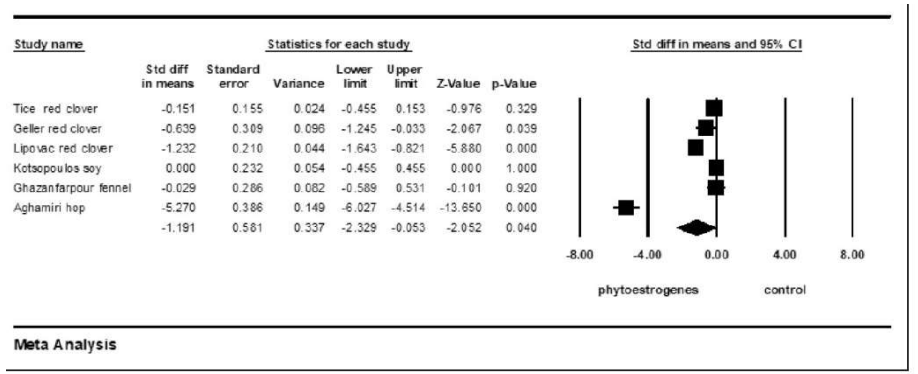
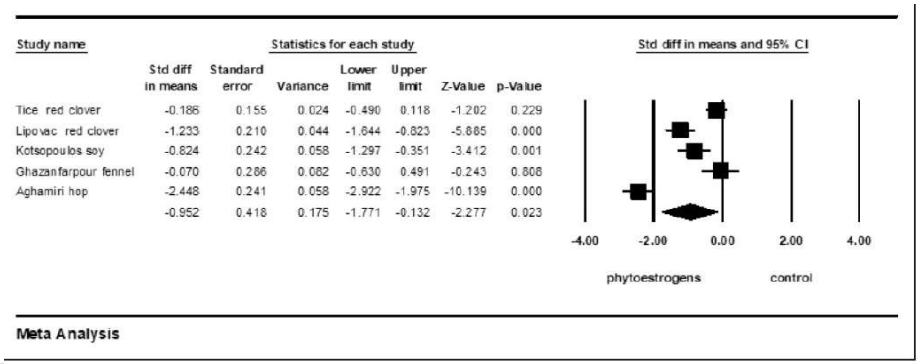
Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Shinjyo 2020	Jenabi 2017	Menopausal symptoms	Valerian root	Vasomotor symptoms	8 weeks	Hot flashes, daily frequency	Higher is worse	60 (NR)	NR	NR	NR*	NR	*Not reported*	Some concerns
	Mirabi 2013	Menopausal symptoms	Valerian root	Vasomotor symptoms	8 weeks	Hot flashes, daily frequency	Higher is worse	68 (NR)	NR	NR	NR*	NR	*Not reported*	Some concerns
	Taavoni 2011	Menopausal symptoms	Valerian root	Sleep quality	4 weeks	PSQI	Higher is worse	100 (NR)	NR	NR	NR*	NR	*Not reported*	Some concerns
	Footnotes:	*No further details provided.												
Ghorbani 2019	Oh 2010	Menopause (1-yr amenorrhea)	Ginseng powder	Sexual function	8 weeks	FSFI	Higher is worse							High
	Chung 2015	Menopausal symptoms	Ginseng powder	Sexual function	8 weeks	FSFI	Higher is worse							High
	Kim 2009*	Menopausal symptoms	Ginseng powder	Sexual function	6 weeks	FSFI	Higher is worse							High
				Quality of life	6 weeks	SF-36	NR							High
	Dongre 2015	Menopausal symptoms	Ginseng powder	Sexual function	8 weeks	FSFI	Higher is worse							Low
	Wiklund 1999	Menopausal symptoms	Ginseng	Quality of life	16 weeks	Women's health questionnaire	NR							Some concerns
				Psychosocial wellbeing	16 weeks	PGWBI	NR							Some concerns
	Footnotes:	*placebo arm double counted.												
	Geller 2009	Menopausal symptoms	Red clover	Anxiety	12 weeks	NR	NR	Data reported in forest plots below						Some concerns
	Charandabi 2013	Menopausal symptoms	Red clover	Psychosocial wellbeing	8 weeks	NR	NR							Some concerns
	Aghamiri 2016	Menopausal symptoms	Red clover	Depression	12 weeks	NR	NR							Some concerns
				Anxiety	12 weeks	NR	NR							
	Rahimi Kian 2017	Menopausal symptoms	Red clover	Psychosocial wellbeing	NR	MenQoL-psychosocial	NR							Some concerns

Figure. 4. Meta-Analysis of Ginseng Interventions for Improving Sexual Function.

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Shahmohammadi 2019	Ghazanfarpour 2018	Menopausal symptoms	Red clover	Depression	12 weeks	NR								Some concerns
				Anxiety	12 weeks	NR								
	Steels 2017	Menopausal symptoms	Red clover	Psychosocial wellbeing	12 weeks	NR								Low
	Shamshad Begum 2016	Menopausal symptoms	Red clover	Depression	13 weeks	NR								Low
				Anxiety	13 weeks	NR								
	Lambert 2017	Menopausal symptoms	Red clover	Psychosocial wellbeing	12 weeks	GCS-psychosocial								Some concerns
	Hidalgo 2005	Menopausal symptoms	Red clover	Depression	12 weeks	KMI-depression								High
	Tice 2003	Menopausal symptoms	Red clover	Depression	12 weeks	NR								Low
				Anxiety	12 weeks	NR								
				Psychosocial wellbeing	12 weeks	NR								
	Ehsanpour 2012	Menopausal symptoms	Red clover	Psychosocial wellbeing	8 weeks	NR								Some concerns
	Lipovac 2012	Menopausal symptoms	Red clover	Depression	12 months	NR								Some concerns
				Anxiety	12 months	NR								
	Shakeri 2015	Menopausal symptoms	Red clover	Psychosocial wellbeing	12 weeks	NR								Low
	Footnotes:	No other data provided.												
	Steels 2017	Menopausal symptoms	Fenugreek (dehusked)	Quality of life	12 weeks	MenQoL	NR	NR	NR	NR	NR	NR	NR	Low
	Rahimi Kian 2017	Menopausal symptoms	Fennel	Quality of life	NR	MenQoL	NR	NR	NR	NR	NR	NR	NR	Some concerns
	Shamshad Begum 2016	Menopausal symptoms	Fenugreek husk	Symptom severity	13 weeks	GCS	NR	NR	NR	NR	NR	NR	NR	Low

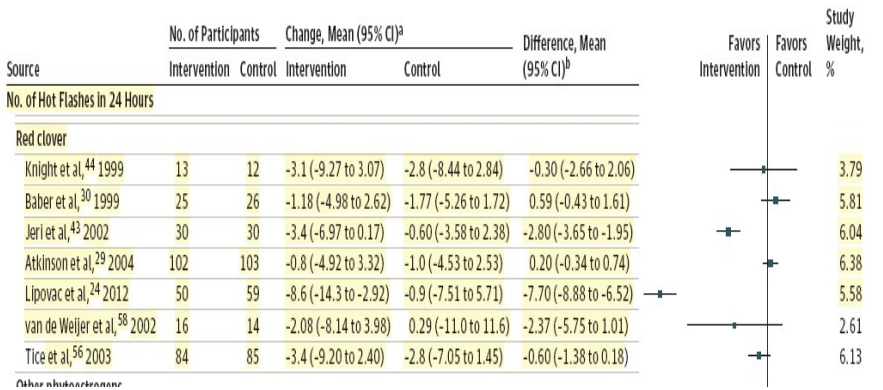


ure 2. Effect of phytoestrogens on anxiety, ■ point estimate; ◆ combined overall effect of treatment. CI: confidence interval.



gure 3. Effect of phytoestrogens on depression, ■ point estimate; ◆ combined overall effect of treatment. CI: confidence interval.

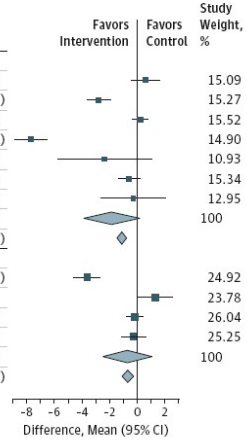
Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Najafi 2018a	Shakeri 2015	Menopausal symptoms	Red clover	Symptom severity	12 weeks	MRS total	NR	NR	NR	NR	NR	NR	NR	Low
	Ehsanpour 2012	Menopausal symptoms	Red clover	Sexual function	8 weeks	MenQoL-sexual function	NR	NR	NR	NR	MD 1.100 (1.776, 3.976)	0.453	*No difference*	Some concerns
	del Giorno 2010	Menopausal symptoms	Red clover	Sexual function	12 months	NR	NR	NR	NR	NR	NR	NR	NR	Low
	Oh 2010	Menopausal symptoms	Red Ginseng	Sexual function	8 weeks	FSFI	NR	NR	NR	NR	NR	NR	NR	High
	Tice 2003	Menopausal symptoms	Red clover 41 mg	Symptom severity	12 weeks	GCS	NR	NR	NR	NR	MD 0.200 (1.089, 0.689)	0.659	*No difference*	Low
			Red clover 28.6 mg	Symptom severity	12 weeks	GCS	NR	NR	NR	NR	MD -0.087 (0.936, 0.763)	0.763	*No difference*	Low
	Footnotes:													
Franco 2016	Knight 1999	Menopausal symptoms	Red clover	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency								
	Baber 1999	Menopausal symptoms	Red clover	Vasomotor symptoms	90 days	Hot flashes, daily frequency								
	Jeri 2002	Menopausal symptoms	Red clover	Vasomotor symptoms	16 weeks	Hot flashes, daily frequency								
	Atkinson 2004	Menopausal symptoms	Red clover	Vasomotor symptoms	12 months	Hot flashes, daily frequency								
	Lipovac 2012	Menopausal symptoms	Red clover	Vasomotor symptoms	12 months	Hot flashes, daily frequency								
	van de Weijer 2002	Menopausal symptoms	Red clover	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency								
	Tice 2003	Menopausal symptoms	Red clover	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency								
	Charandabi 2013	Menopausal symptoms	Black cohosh	Vasomotor symptoms	8 weeks	Hot flashes, daily frequency								
	Chung 2007	Menopausal symptoms	Black cohosh	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency								



Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Franco 2016 cont'd	Frei-Kleiner 2005	Menopausal symptoms	Black cohosh	Vasomotor symptoms	12 weeks	Hot flashes, daily frequency								
	Newton 2006	Menopausal symptoms	Black cohosh	Vasomotor symptoms	48 weeks	Hot flashes, daily frequency								
	Pockaj 2006	Menopausal symptoms	Black cohosh	Vasomotor symptoms	8 weeks	Hot flashes, daily frequency								
	Shahnazi 2013	Menopausal symptoms	Black cohosh	Vasomotor symptoms	8 weeks	Hot flashes, daily frequency								
	Jiang 2015	Menopausal symptoms	Black cohosh	Vasomotor symptoms	24 weeks	Hot flashes, daily frequency								
	Abdali 2010	Menopausal symptoms	Black cohosh	Vasomotor symptoms	8 weeks	Hot flashes, daily frequency								
	van Die 2009	Menopausal symptoms	Black cohosh	Vasomotor symptoms	17 weeks	Hot flashes, daily frequency								
	Footnotes:													

Figure 3. Meta-analysis of Randomized Clinical Trials Assessing the Associations Between Use of Red Clover and Black Cohosh and Number of Daily Hot Flashes

Source	Plant-Based Therapy Dosage, mg	No. of Participants		Change, Mean (95% CI)[a]		Difference, Mean (95% CI)[b]	Study Weight, %
		Intervention	Control	Intervention	Control		
Red Clover							
Baber et al,[30] 1999	40	25	26	-1.18 (-4.98 to 2.62)	-1.77 (-5.26 to 1.72)	0.59 (-0.43 to 1.61)	15.09
Jeri et al,[43] 2002	40	30	30	-3.4 (-6.97 to 0.17)	-0.60 (-3.58 to 2.38)	-2.80 (-3.65 to -1.95)	15.27
Atkinson et al,[29] 2004	40	102	103	-0.8 (-4.92 to 3.32)	-1.0 (-4.53 to 2.53)	0.20 (-0.34 to 0.74)	15.52
Lipovac et al,[24] 2012	80	50	59	-8.6 (-14.3 to -2.92)	-0.9 (-7.51 to 5.71)	-7.70 (-8.88 to -6.52)	14.90
van de Weijer et al,[58] 2002	80	16	14	-2.08 (-8.14 to 3.98)	0.29 (-11.0 to 11.6)	-2.37 (-5.75 to 1.01)	10.93
Tice et al,[56] 2003	82	84	85	-3.4 (-9.20 to 2.40)	-2.8 (-7.05 to 1.45)	-0.60 (-1.38 to 0.18)	15.34
Knight et al,[44] 1999	160	13	12	-3.1 (-9.27 to 3.07)	-2.8 (-8.44 to 2.84)	-0.30 (-2.66 to 2.06)	12.95
Random effects						-1.84 (-3.87 to 0.19)	100
Fixed effects						-1.12 (-1.46 to -0.77)	
Black Cohosh							
Shahnazi et al,[71] 2013	6.5	42	42	-4.6 (-3.65 to -3.25)	-1.19 (-2.74 to 0.36)	-3.64 (-4.61 to -2.67)	24.92
Pockaj et al,[70] 2006	40	66	65	NR	NR	1.32 (0.07 to 2.57)	23.78
Frei-Kleiner et al,[68] 2005	42	81	41	1.66 (-1.65 to 4.97)	1.85 (-1.33 to 5.03)	-0.19 (-0.81 to 0.43)	26.04
Newton et al,[19] 2006	160	80	84	NR	NR	-0.28 (-1.16 to 0.60)	25.25
Random effects						-0.71 (-2.51 to 1.08)	100
Fixed effects						-0.69 (-1.12 to -0.27)	



Assessment of heterogeneity: red clover and number of hot flashes in 24 hours, I^2 = 97% (95% CI, 95%-98%; P < .001); black cohosh and number of hot flashes: I^2 = 60% (95% CI, 0%-89%; P = .08). Sizes of data markers are proportional to the inverse of the variance of the effect estimate. NR indicates not reported.

[a] Mean change in the number of hot flashes in 24 hours from randomization to the end of study.

[b] Mean difference of changes in the number of hot flashes in 24 hours between treatment groups.

WHM vs inactive control (no intervention, waitlist, usual care)														
No RCTs reporting this comparison were identified by the eligible reviews (see below)														
WHM vs active control														
	Bai 2007	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported
	Chen 2013	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported
	Chen 2014	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Castelo-Branco 2021*	Huang 2013	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported
	Liske 2002	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported
	Nappi 2005	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported
	Sun 2012	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported
	Wang 2019	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported
	Xi 2014	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported
	Zhang 2015	Menopausal symptoms	black cohosh extract vs NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	*Not reported*	Not reported
	Footnotes:	*The review authors identify 10 RCTs comparing black cohosh extract with other interventions (such as hormone therapy, vitamins/minerals, or antidepressants) but details about these studies were not provided.												
Firoozeei 2021	Kamalifard 2017	Menopausal symptoms	Oral lavender vs oral bitter orange	Depression	8 weeks	BDI	Higher is worse	156 (NR)	NR	NR	NR*	NR	*No difference*	Low
	Footnotes:	*No further details provided.												

Abbreviations: C, Comparator; Confidence interval; I, intervention; NR, not reported; PGWBI, Psychological General Well-Being Index; RoB, risk of bias; SD, standard deviation

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
WHM vs placebo														
Ghaderi 2020*	Jafarnia 2017	GAD	Saffron	Anxiety	4 weeks	HAM-A	higher is worse	20/20	NR	NR	WMD: -1.90 (-3.44, -0.36)	NR	*Favours intervention*	Some concerns
	Mazidi 2016	Symptoms of Anxiety	Saffron	Anxiety	12 weeks	Beck Anxiety Inventory	higher is worse	30/24	NR	NR	WMD: -3.19 (-4.51, -1.87)	NR	*Favours intervention*	Some concerns
				Depression	12 weeks	Beck Depression Inventory	higher is worse	30/24	NR	NR	WMD: -2.34 (-4.32, -0.36)	NR	*Favours intervention*	Some concerns
	Footnotes:	*the authors focus was effect so saffron on variety of outcomes. Only data from studies on people with anxiety (+/- depression) included here.												
Janda 2020*	Akhondzadeh 2001	GAD	Passiflora	Anxiety	4 weeks	HAM-A	higher is worse	36 (18/18)	5.5 (0.75)	5.1 (1.28)	NR	NR	*Favours intervention*	Some concerns
	Footnotes:	*the authors focus was the effect of passiflora on neuropsychiatric conditions. Only data from studies on people with anxiety (+/- depression) included here.												
Sayad 2020		Authors presented a network meta-analysis. Individual study data not provided.												
	Footnotes:	* All studies included and reported by Donelli 2019												
Shinjyo 2020*	Andreatini 2002	GAD	Valerian extract	Anxiety	4 weeks	HAM-A	higher is worse	12/12	Total HAM-A score: significant reduction in all 3 groups. (No significant difference among the 3 groups.) No				*No difference*	Some concerns
	Footnotes:	*the authors focus was the effect of Valerian extrant on sleep conditions and related disorder. Only data from studies on people with anxiety (+/- depression) included here.												
Donelli 2019*	Kasper 2016	Symptoms of Anxiety	Lavender oil (Silexan)	Anxiety	NR	HAM-A	higher is worse	159/156	-10.8 (9.6)	-8.4 (8.9)	MD: -2.4 (-4.44, -0.36)	NR	*Favours intervention*	High
	Kasper 2017	Symptoms of Anxiety	Lavender oil (Silexan)	Anxiety	NR	HAM-A	higher is worse	103/102	-11..6 (8.1)	-11.4 (8)	MD: -0.20 (-2.40, 2.00)	NR	*No difference*	High
	Kasper 2010	Symptoms of Anxiety	Lavender oil (Silexan)	Anxiety	NR	HAM-A	higher is worse	104/108	-16 (8.3)	-9.5 (9.1)	MD: -6.5 (-8.84, -4.16)	NR	*Favours intervention*	Low
	Kasper 2014	GAD	Lavender oil (Silexan)	Anxiety	NR	HAM-A	higher is worse	135/136	-12.8 (8.7)	-9.5 (9)	MD: -3.30 (-5.41, -1.19)	NR	*Favours intervention*	Some concerns
	Kasper 2015	Symptoms of Anxiety	Lavender oil (Silexan)	Anxiety	NR	HAM-A	higher is worse	86/84	-11.8 (7.1)	-9.6 (9.1)	MD: -2.20 (-4.66, 0.26)	NR	*No difference*	High
	Footnotes:	*80 mg results reported												

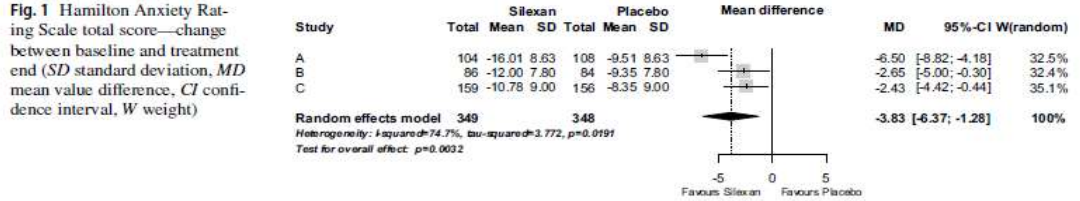
Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Hieu 2019	Amsterdam 2009 (Keefe 2016)*	GAD	German chamomile	Anxiety	8 weeks	HAM-A	higher is worse	28/29	-8.2 (5.3329)	-4.80 (6.0841)	MD: -3.40 (-6.37, -0.43)	NR	*Favours intervention*	High
	Mao 2016	GAD	German chamomile	Anxiety	8 weeks	Beck Anxiety Inventory	higher is worse	46/47	1.0 (4.5)	1.5 (4.7)	MD: -0.50 (-2.37, 1.37)	NR	*No difference*	High
	Footnotes:	*Amsterdam et al. found no statistically significant effect of chamomile on BAI, PGWB index, and CGI/S rating (P > 0.05; Amsterdam et al., 2009).												
Marx 2019 ^	Lopresti 2018	GAD (children)	Saffron	Anxiety	6 weeks	RCADS	NR	38/37	NR	NR	Hedges's g: 0.739	0.002	*Favours intervention*	Not reported
				Depression	6 weeks	RCADS	NR	38/37	NR	NR	Hedges's g: 0.569	0.015	*Favours intervention*	Not reported
	Footnotes:	^only RCT data not already extracted above reported here.												
Moler 2019	Study A: Kasper 2010	Symptoms of Anxiety	Lavender oil (Silexan)											
	Study B: Kasper 2015	Symptoms of Anxiety	Lavender oil (Silexan)											
	Study B: Kasper 2016	Symptoms of Anxiety	Lavender oil (Silexan)											
	Footnotes:													
c 2018	Volz 1997	GAD	Kava Kava	Anxiety	24 weeks	HAM-A	higher is worse	52/48	9.7 (9.9)	15.2 (9.6)	MD: -5.50 (-9.33, -1.67)	0.005	*Favours intervention*	Some concerns
	Malsch 2001	GAD	Kava Kava	Anxiety	5 weeks	HAM-A#	higher is worse	20/20	NR	NR	NR	NR	*Not reported*	Some concerns
	Connor 2002	GAD	Kava Kava	Anxiety	3 weeks	HAM-A	higher is worse	17/18	14.2 (8.3)	10.3 (4.4)	MD: 3.90 (-0.47, 8.27)	0.08	*Favours comparator*	Some concerns
	Sarris 2013	GAD	Kava Kava	Anxiety	6 weeks	HAM-A	higher is worse	27/31	14 (7)	15.3 (6.2)	MD: -1.23 (-4.63, 2.17)	0.478	*No difference*	Low
	Kasper 2014	GAD	Lavender oil (Silexan) 80mg/day	Anxiety	10 weeks	HAM-A	higher is worse	135/135	-12.8 (8.7)	-9.5 (9)	MD: -3.30 (-5.41, -1.2)	0.002	*Favours intervention*	Some concerns

Fig. 1 Hamilton Anxiety Rating Scale total score—change between baseline and treatment end (*SD* standard deviation, *MD* mean value difference, *CI* confidence interval, *W* weight)

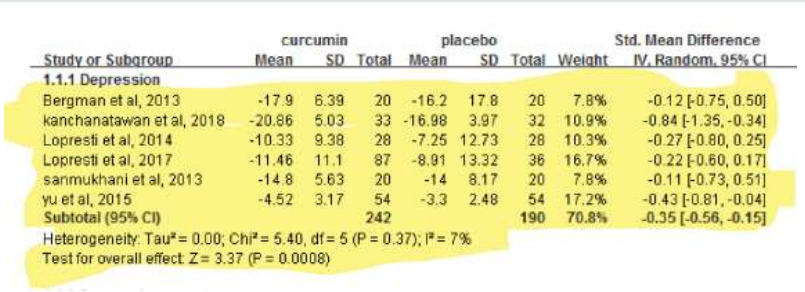
Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Bari	Kasper 2014	GAD	Lavender oil (Silexan) 160mg/day	Anxiety	10 weeks	HAM-A	higher is worse	121/135	-14.1 (9.3)	-9.5 (9)	MD: -4.6 (-6.9, -2.3)	<0.001	*Favours intervention*	Some concerns
	Amsterdam 2009	GAD	German chamomile	Anxiety	8 weeks	HAM-A	higher is worse	28/29	NR	NR	MD: -3.2 (-6.3, -0.45)	0.047	*Favours intervention*	High
	Mao 2016	GAD	German chamomile	Anxiety	8 weeks	CGI-S	higher is worse	46/47	NR	NR	NR	NR	*No difference*	High
	Andreatini 2002	GAD	Valerian extract	Anxiety	4 weeks	HAM-A	higher is worse	12/12	14.6 (9.8)	16.0 (6.1)	NR	NR	*No difference*	Some concerns
	Footnotes:	#trialists report end of treatment responders and suggest and effect in favour of intervention (12/20 vs 4/20, OR 6.00; p=0.013)												
Ooi 2018 ^	Connor 2006	GAD	Kava Kava	Anxiety	Study discontinued									
	Savage 2015	GAD	Kava Kava	Anxiety	Protocol only. Study not published NCT02219880									
	Footnotes:	^only RCT data not extracted elsewhere reported here.												
Smith 2018^	Sarris 2009	Symptoms of Anxiety	Kava Kava	Anxiety	3 weeks	HAM-A	higher is worse	29/18	11.26 (4.47)	19.5 (7.26)	NR	0.0001	*Favours intervention*	Not reported
	Geier 2004	Symptoms of Anxiety	Kava Kava	Anxiety	7 weeks	HAM-A	higher is worse	25/25	14.8 (4.3)	16.8 (3.55)	NR	0.1	*No difference*	Not reported
	Lehrl 2004	Symptoms of Anxiety	Kava Kava	Anxiety	7 weeks	HAM-A	higher is worse	34/23	median 11	median 14	Median Diff 3.00 (IQR: 7 to -4)	0.1	*No difference*	Not reported
	Gastpar 2003	Symptoms of Anxiety	Kava Kava	Anxiety	7 weeks	ASI	higher is worse	71/70	39 (2.35)	40.6 (2.3)	NR	NS	*No difference*	Not reported
	Footnotes:	^only RCT data not already extracted above reported here.												
Brondino 2013^	Woelk 2007	GAD	Ginkgo	Anxiety	4 weeks	HAM-A	higher is worse	27/25/30	Authors note response rates (50% reduction in score) of 44% vs 31% vs 22% in high-dose,			significant	*Favours intervention*	Low
	Footnotes:	^ Study data not adequately reported. SR authors provide narrative summary only.												
opresti 2021^	Lopresti 2019	Symptoms of Anxiety	Ashwganda	Anxiety	8 weeks	HAM-A	higher is worse	40/20	NR	NR	NR	NR	*Not reported*	Low
	Kyati 2013	GAD	Ashwganda	Anxiety	8 weeks	HAM-A	higher is worse	44/42	NR	NR	NR	NR	*Not reported*	Low
	Auddy 2008	Symptoms of Anxiety	Ashwganda	Anxiety	8 weeks	HAM-A	higher is worse	100/30	NR	NR	NR	NR	*Not reported*	Low

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Lof	Andrade 2000	Symptoms of Anxiety	Ashwganda	Anxiety	8 weeks	HAM-A	higher is worse	20/19	NR	NR	NR	NR	*Not reported*	Low
	Footnotes:	^ Study data not adequately reported. SR authors provide narrative summary only.												
WHM vs inactive control (no intervention, waitlist, usual care)														
No RCTs reporting this comparison were identified by the eligible reviews														
WHM vs active control														
Shinjyo 2020	Andreatini 2002	GAD	Valerian extract vs diazepam	Anxiety	4 weeks	HAM-A	higher is worse	12/12	Total HAM-A score: significant reduction in all 3 groups. (No significant difference among the 3 groups.) No other data provided.				*No difference*	Some concerns
	Footnotes:	*the authors focus was the effect of Valerian extrant on sleep conditions and related disorder. Only data from studies on people with anxiety (+/- depression) included here.												
Donelli 2019	Woelk 2010	Symptoms of Anxiety	Lavender oil (Silexan) vs diazepam	Anxiety	6 weeks	Zung SAS score	higher is worse	36/33	-14.8 (11.4)	-14.4 (8.5)	MD: -0.40 (-5.12, 4.32)	NR	*No difference*	High
	Footnotes:													
Hieu 2019	Amsterdam 2009 (Keefe 2016)	GAD	German chamomile vs diazepam	Anxiety	8 weeks	HAM-A	higher is worse	28/29	-8.2 (5.3329)	NR	NR	NR	*Not reported*	High
	Mao 2016	GAD	German chamomile vs diazepam	Anxiety	8 weeks	Beck Anxiety Inventory	higher is worse	46/47	1.0 (4.5)	NR	NR	NR	*Not reported*	High
	Footnotes:													
	Boerner 2003	GAD	Kava Kava vs buspirone OR opipramol	Anxiety	8 weeks	HAM-A	higher is worse	43/43	8.4 (7.4)	7.9 (7.6)	MD: 0.56 (-2.25, 3.66)	0.693	*No difference*	Low
	Woelk 2010	Symptoms of Anxiety	Lavender oil (Silexan) vs diazepam	Anxiety	6 weeks	HAM-A	change from baseline	40/37	-11.3 (6.7)	-11.6 (6.6)	MD: -0.30 (-2.7, 3.3)	0.844	*No difference*	High

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	direction of effect	RoB
Baric 2018	Kasper 2014	GAD	Lavender oil (Silexan) 80mg/day vs paroxetine 20 mg/day	Anxiety	10 weeks	HAM-A	higher is worse	135/132	-12.8 (8.7)	-11.3 (8)	NR	NR	*Not reported*	Some concerns
			Lavender oil (Silexan) 160mg/day vs paroxetine 20 mg/day	Anxiety	10 weeks	HAM-A	higher is worse	121/132	-14.1 (9.3)	-11.3 (8)	NR	NR	*Not reported*	Some concerns
	Andreatini 2002	GAD	Valerian extract vs diazepam	Anxiety	4 weeks	HAM-A	higher is worse	12/12	14.6 (9.8)	14.2 (6.3)	NR	NR	*No difference*	Some concerns
	Footnotes:													

Abbreviations: C, Comparator; Confidence interval; CGI-S, Clinical global impressions - symptoms; GAD, Generalised Anxiety Disorder; HAM-A, Hamilton Anxiety Rating Scale; I, intervention; NR, not reported; RCADS, Child Anxiety & Depression Scale-revised; RoB, risk of bias; SAS, Zung Anxiety Self-rating Scale; SD, standard deviation

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
WHM (other than St John's wort) # vs placebo														
Wang 2021	Kanchanatawan 2018	Major Depressive Disorder	Curcumin	Depression	12 weeks	MADRS	higher is worse							Low
	Lopresti 2017	Major Depressive Disorder	Curcumin	Depression	8 weeks	IDR-SR30	higher is worse							Low
	Yu 2015	Major Depressive Disorder	Curcumin	Depression	6 weeks	HAM-D	higher is worse							High
	Lopresti 2014	Major Depressive Disorder	Curcumin	Depression	12 weeks	IDS-SR30	higher is worse							Low
	Sanmukhani 2014	Major Depressive Disorder	Curcumin	Depression	6 weeks	HAM-D	higher is worse							High
	Bergman 2013	Major Depressive Disorder	Curcumin	Depression	6 weeks	HAM-D	higher is worse							Some concerns
	Footnotes:													
Dai 2020	Tabeshpour 2017	Postpartum depression	Saffron	Depression	8 weeks	BDI	higher is worse	30/20	NR	NR	MD: -7.60 (-10.22, -4.98)	NR	*Favours intervention*	Some concerns
	Moshiri 2006	Mild-Moderate Depression	Saffron	Depression	6 weeks	HAM-D	higher is worse	20/30	NR	NR	MD: -8.96 (-12..12, -5.80)	NR	*Favours intervention*	Low
	Akhondzadeh 2005	Mild-Moderate Depression	Saffron	Depression	6 weeks	HAM-D	higher is worse	20/20	NR	NR	MD: -7.10 (-10.01, -4.19)	NR	*Favours intervention*	Low
	Footnotes:													

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Fusar-Poll 2020	Kanchanatawan 2018	Major Depressive Disorder	Curcumin	Depression	12 weeks	MADRS	higher is worse	33/32	Available data reported below					Some concerns
				Anxiety	12 weeks	HAM-A	higher is worse	33/32						Some concerns
	Lopresti 2017	Major Depressive Disorder	Curcumin	Depression	8 weeks	IDR-SR30	higher is worse	28/28						Low
				Anxiety	8 weeks	STAI	higher is worse							Low
	Yu 2015	Major Depressive Disorder	Curcumin	Depression	6 weeks	HAM-D	higher is worse	54/54						Some concerns
	Lopresti 2014	Major Depressive Disorder	Curcumin	Depression	12 weeks	IDS-SR30	higher is worse	33/36						Low
				Anxiety	12 weeks	STAI	higher is worse							Low
	Sanmukhani 2014	Major Depressive Disorder	Curcumin	Depression	6 weeks	HAM-D	higher is worse	20/20						High
				Improvement	6 weeks	CGI	NR	20/20						High
	Bergman 2013	Major Depressive Disorder	Curcumin	Depression	6 weeks	HAM-D	higher is worse	20/20						High
				Improvement	6 weeks	CGI	NR	20/20						High
	Panahi 2015	Major Depressive Disorder	Curcumin	Depression	6 weeks	BDI	higher is worse	61/50						High
				Anxiety	6 weeks	HADS-A	higher is worse							High

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Fusar-Poll 2020 cont'd	Footnotes:	a) Depression Statistics for each study Study name | Hedges's g | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value Bergman et al. 2013 | -0,295 | 0,316 | 0,100 | -0,914 | 0,323 | -0,936 | 0,349 Esmaily et al. 2015 | -0,202 | 0,356 | 0,127 | -0,900 | 0,497 | -0,566 | 0,571 Kanchanatawan et al. 2018 | -0,415 | 0,256 | 0,065 | -0,916 | 0,086 | -1,622 | 0,105 Lopresti et al. 2014 | -0,350 | 0,266 | 0,071 | -0,870 | 0,171 | -1,317 | 0,188 Lopresti et al. 2017 (low dosage) | -1,825 | 0,376 | 0,142 | -2,563 | -1,087 | -4,847 | 0,000 Lopresti et al. 2017 (high dosage) | -1,718 | 0,353 | 0,125 | -2,410 | -1,026 | -4,865 | 0,000 Panahai et al. 2015 | -0,951 | 0,200 | 0,040 | -1,343 | -0,559 | -4,757 | 0,000 Sanmukhani et al. 2013 | -0,185 | 0,341 | 0,116 | -0,853 | 0,482 | -0,544 | 0,586 Setiawati et al. 2017 | -2,124 | 0,684 | 0,468 | -3,464 | -0,783 | -3,105 | 0,002 Yu et al. 2015 | -0,425 | 0,201 | 0,040 | -0,819 | -0,032 | -2,119 | 0,034 -0,754 | 0,184 | 0,034 | -1,115 | -0,394 | -4,099 | 0,000 b) Anxiety Statistics for each study Study name | Hedges's g | Standard error | Variance | Lower limit | Upper limit | Z-Value | p-Value Esmaily et al. 2015 | -2,532 | 0,483 | 0,233 | -3,479 | -1,586 | -5,245 | 0,000 Lopresti et al. 2014 | -0,264 | 0,265 | 0,070 | -0,783 | 0,255 | -0,998 | 0,318 Lopresti et al. 2017 (low dosage) | -4,713 | 0,610 | 0,372 | -5,908 | -3,517 | -7,727 | 0,000 Lopresti et al. 2017 (high dosage) | -4,401 | 0,551 | 0,303 | -5,481 | -3,322 | -7,994 | 0,000 Panahi et al. 2015 | -1,548 | 0,216 | 0,047 | -1,971 | -1,124 | -7,164 | 0,000 -2,617 | 0,736 | 0,541 | -4,059 | -1,175 | -3,557 | 0,000 Figure 3. Meta-analyses of the effect of curcumin in people with depressive disorders on the following out													
Ghaderi 2020^	Talaei 2015	Major Depressive Disorder	Saffron	Depression	4 weeks	BDI	higher is worse	20/20	NR	NR	MD -11.50 (-13.66, -9.34)	NR	*Favours intervention*	Not reported
	Sahraian 2016	Major Depressive Disorder	Saffron	Depression	4 weeks	BDI	higher is worse	11/19	NR	NR	MD 1.03 (-0.63, 2.69)	NR	*Favours comparator*	Not reported
	Jelodar 2018	Major Depressive Disorder	Saffron	Depression	4 weeks	BDI	higher is worse	20/20	NR	NR	MD -3..45 (-3.72, -3.18)	NR	*Favours intervention*	Not reported
	Kell 2017	Low mood	Saffron 18 mg/day	Sleep quality	4 weeks	PSQI	higher is worse	16/37	NR	NR	MD -1.45 (-3.12, 0.22)	NR	*Favours intervention*	Not reported
			Saffron 28 mg/day	Sleep quality	4 weeks	PSQI	higher is worse	17/39	NR	NR	MD -1.87 (-3.42, -0.32)	NR	*Favours intervention*	Not reported
	Footnotes:	^only RCT data not already extracted above reported here.												

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Khaksarian 2019^	Akhondzadeh Basti 2007	Mild-Moderate Depression	Saffron	Depression	6 weeks	HAM-D	higher is worse	19/25	-12.18 (3.72)	-13.45 (4.84)	SMD 0.28 (-0.32, 0.88)	NR	*No difference*	Some concerns
	Footnotes:	^only RCT data not already extracted above reported here.												
Marx 2019^	Kashani 2013	Major Depressive Disorder	Saffron	Depression	4 weeks	HAM-D	higher is worse	34 (NR)	NR	NR	Hedges's g 0.363 (-0.299, 1.025)	0.283	*Favours intervention*	Some concerns
	Modabberni a 2012	Major Depressive Disorder	Saffron	Depression	4 weeks	HAM-D	higher is worse	30 (NR)	NR	NR	Hedges's g 0.485 (-0.222, 1.192)	0.179	*Favours intervention*	Some concerns
	Footnotes:	^only RCT data not already extracted above reported here.												
Toth 2019	No additional studies found. All available data already extracted													
Yang 2019	Moshiri 2006	Mild-Moderate Depression	Saffron	Depression	6 weeks	HAM-D								
	Akhondzadeh 2005	Mild-Moderate Depression	Saffron	Depression	6 weeks	HAM-D								
	Footnotes:	* study by Mazidi 2016 is in people with anxiety and included previously.												
WHM (other than St John's wort)# vs inactive control (no intervention, waitlist, usual care)														
No RCTs reporting this comparison were identified by the eligible reviews														
WHM (other than St John's wort) # vs active control														
Firoozeel 2021	Araj-Khodaei 2020	Mild-Moderate Depression	Lavender (oral) vs fluoxetine	Depression	8 weeks	HAM-D	higher is worse	NR	NR	NR	SMD 0.57 (-0.12, 1.26)	0.877	*Favours comparator*	High
	Footnotes:													
	Ghajar 2017	Major Depression	Saffron vs citalopram	Depression	6 weeks	HAM-D	higher is worse	30/20	NR	NR	MD 1.14 (-1.93, 4.21)	NR	*Favours comparator*	Low

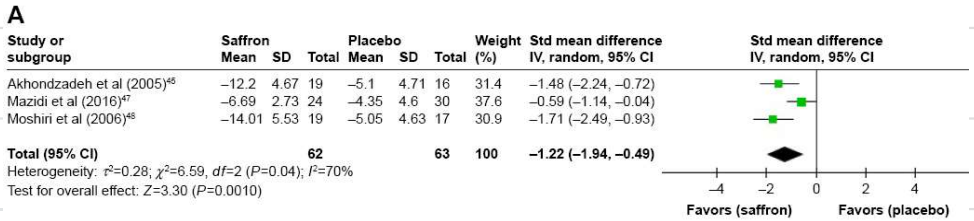


Figure 3 Meta-analyses of primary outcomes: (**A**) improvement of depression symptoms compared with placebo;

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Dai 2020	Kashani 2016	Postpartum depression	Saffron vs fluoxetine	Depression	6 weeks	HAM-D	higher is worse	32/17	NR	NR	MD 0.21 (-0.69, 1.11)	NR	*Favours comparator*	Some concerns
	Akhondzadeh Basti 2007	Mild-Moderate Depression	Saffron vs fluoxetine	Depression	8 weeks	HAM-D	higher is worse	20/19	NR	NR	MD 1.5 (-1.30, 4.30)	NR	*Favours comparator*	Some concerns
	Noorbala 2005	Mild-Moderate Depression	Saffron vs fluoxetine	Depression	6 weeks	HAM-D	higher is worse	20/20	NR	NR	MD 2.80 (-0.53, 6.13)	NR	*Favours comparator*	Low
	Footnotes:													
Marx 2019^	Akhondzadeh 2004	Major Depression	Saffron vs Imipramine	Depression	6 weeks	HAM-D	higher is worse	30 (NR)	NR	NR	Hedges's g 0.606	0.096	*Favours intervention*	Some concerns
	Footnotes:	^only RCT data not already extracted above reported here.												
Khaksarian 2019^	Akhondzadeh Basti 2007	Mild-Moderate Depression	Saffron vs fluoxetine	Depression	8 weeks	HAM-D	higher is worse	20/20	-12.00 (4.10)	-13.50 (4.91)	SMD 0.33 (-0.30, 0.95)	NR	*Favours comparator*	Some concerns
	Noorbala 2005	Mild-Moderate Depression	Saffron vs fluoxetine	Depression	6 weeks	HAM-D	higher is worse	20/20	-12.20 (4.67)	-15.00 (5.88)	SMD 0.52 (-0.11, 1.15)	NR	*Favours comparator*	Low
	Kashani 2016	Postpartum depression	Saffron vs fluoxetine	Depression	6 weeks	HAM-D	higher is worse	32/32	7.50 (1.97)	7.71 (1.69)	SMD -0.11 (-0.60, 0.38)	NR	*No difference*	Some concerns
	Footnotes:													
Yang 2019	Akhondzadeh Basti 2007	Mild-Moderate Depression	Saffron vs fluoxetine	Depression	8 weeks	HAM-D								
	Akhondzadeh 2004	Major Depression	Saffron vs Imipramine	Depression	6 weeks	HAM-D								
	Ghajar 2017	Major Depression	Saffron vs citalopram	Depression	6 weeks	HAM-D								
	Noorbala 2005	Mild-Moderate Depression	Saffron vs fluoxetine	Depression	6 weeks	HAM-D								
	Footnotes:													

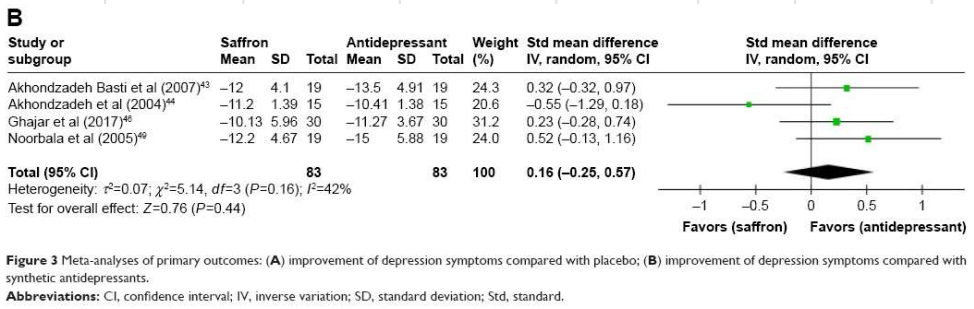


Figure 3 Meta-analyses of primary outcomes: (**A**) improvement of depression symptoms compared with placebo; (**B**) improvement of depression symptoms compared with synthetic antidepressants.
Abbreviations: CI, confidence interval; IV, inverse variation; SD, standard deviation; Std, standard.

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Toth 2019^	Ghajar 2017	Major Depression	Saffron vs citalopram	Depression	6 weeks	HAM-D	higher is worse	66 (NR)	10.13 (SEM 1.09)	11.27 (SEM 0.67)	NR	NR	*Not reported*	Low
				Anxiety	6 weeks	HAM-A	higher is worse	66 (NR)	NR	NR	NR	NR	*Not reported*	Low
	Footnotes:	^only RCT data not already extracted above reported here.												

Abbreviations: C, Comparator; Confidence interval; I, intervention; NR, not reported; RoB, risk of bias; SD, standard deviation
a comprehensive reivew of St John's wort provided by Apaydin 2016 and not included here.

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
WHM vs placebo														
Lopresti 2021	Langade 2019	Insomnia	Ashwaghanda	Sleep quality	10 weeks	PSQI	higher is worse	60 (40/20)	NR	NR	NR	<0.05	*Favours intervention*	Low
				Symptoms of anxiety		HAM-A	higher is worse		NR	NR	NR	<0.05	*Favours intervention*	Low
Shinjyo 2020	Taibi 2009	Insomnia	Valerian	Sleep quality	2 weeks	self-rated, measure not described; assumed VAS (higher is better)		32*	NR	NR	Hedges's g 0.11 (-0.59, 0.82)	NR	*No difference*	Low
	Oxman 2007	Insomnia	Valerian	Sleep quality	2 weeks	self-rated, measure not described; assumed VAS (higher is better)		405	NR	NR	Hedges's g 0.22 (0.03, 0.42)	0.04	*Favours intervention*	Low
	Donath 2000	Insomnia	Valerian	Sleep quality	2 weeks	VAS	higher is worse	32*	NR	NR	Hedges's g -0.11 (-0.82, 0.59)	NR	*No difference*	High
	Coxeter 2003	Insomnia	Valerian	Sleep quality	3 weeks	VAS	higher is worse	24	NR	NR	NR	NR	*No difference*	Low
	Jacobs 2005	Insomnia	Kava	Insomnia severity	4 weeks	ISI	change from baseline^	(121/135)	NR	8.3	NR	NR	*Not reported*	Low
				Symptoms of anxiety		STAI	change from baseline^		11.8 (12.3)	14.4 (12.9)^	MD 2.7 (-0.8, 6.2)	NR	*Not reported*	Low
			Valerian	Insomnia severity	4 weeks	ISI	change from baseline^	222 (135/135)	NR	8.3	Hedges's g -0.06	NR	*No difference*	Low
				Symptoms of anxiety		STAI	change from baseline^		11.9 (11.9)	14.4 (12.9)^	Hedges's g -0.20 (-0.47, 0.06)	NR	*No difference*	Low
	Taavoni 2011	Insomnia	Valerian	Sleep quality	4 weeks	PSQI	higher is worse	100	NR	NR	Hedges's g 1.28 (0.85, 1.72)	NR	*Favours intervention*	Some concerns
	Morin 2005	Insomnia	Combination (valerian + hops)	HRQoL	4 weeks	NR	higher is best	184	NR	NR	NR	NR	*Favours intervention*	Some concerns
				Sleep quality		self-rated, measure not described; assumed VAS (higher is better)			NR	NR	NR	NR	*No difference*	Some concerns

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
	Footnotes:	*Not clear, but N suggests results are after crossover (N=16 x 2) ^ retrieved from primary study												
Hieu 2019	Zick 2011	Insomnia	Chamomile	Insomnia severity	4 weeks	ISI	higher is worse	34 (17/17)	NR	NR	MD 0.3 (-2.79, 3.39)	0.85	*No difference*	Some concerns
				Sleep quality		PSQI	higher is worse		7.5 (3.3)	7.10 (2.7)	SMD 0.13 (-0.54, 0.8)	NR	*No difference*	Some concerns
				Symptoms of depression		BDI	higher is worse		NR	NR	NR	NR	*Not reported*	Some concerns
				Fatigue		FSS	higher is worse		NR	NR	NR	NR	*Not reported*	Some concerns
				Symptoms of anxiety		STAI	higher is worse		35.5 (11)	40.8 (15.5)	SMD -0.39 (-1.06, 0.29)	NR	*No difference*	Some concerns
	Footnotes:													
Leach 2015	Jacobs 2005	Insomnia	Valerian	Insomnia severity	4 weeks	ISI	change from baseline^	222	NR	8.3	MD 0.4 (-1.3, 2.1)	NR	*No difference*	Some concerns
			Kava	Insomnia severity		ISI	change from baseline^	391	NR	8.3	MD 0.2 (-1.6, 1.9)	NR	*No difference*	Some concerns
	Zick 2011	Insomnia	Chamomile	Sleep quality	6 months	VAS	higher is worse	34	NR	NR	MD -0.4 (-1.07, 0.29)	0.26	*No difference*	Some concerns
				Insomnia severity		ISI	higher is worse		NR	NR	MD 0.3 (-2.79, 3.39)	0.85	*No difference*	Some concerns
	Coxeter 2003	Insomnia	Valerian	Sleep quality	3 weeks	self-rated, measure not described; assumed VAS		24	NR	NR	NR	NR	*Not reported*	Some concerns
	Taibi 2009	Insomnia	Valerian	Sleep quality	2 weeks	self-rated, measure not described; assumed VAS		16 (8/8)	5.9 (1.8)	6.4 (1.4)	SMD -0.29 (-1.28, 0.69)	NR	*No difference*	Low
	Taavoni 2011	Insomnia	Valerian	Sleep quality	4 weeks	PSQI	higher is worse	100 (50/50)	-6.02 (2.6)	-9.4 (3.9)	SMD 1.01 (0.59, 1.43)	NR	*Favours intervention*	Some concerns
	Oxman 2007	Insomnia	Valerian	Sleep quality	2 weeks	self-rated, measure not described; assumed VAS		405 (202/203)	4.06 (1.52)	4.08 (1.29)	SMD -0.01 (-0.21, 0.18)	NR	*No difference*	Some concerns
	Footnotes:													

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Fernández-San-Martín 2010	Leathwood 1985	Sleep problems	Valerian	Sleep quality	18 days	VAS	higher is bétter	6	NR	NR	SMD 0.25 (-0.89, 1.39)	NR	*No difference*	Low
	Vorbach 1996	Sleep problems	Valerian	Sleep quality	4 weeks	short form B	NR	121	NR	NR	NR	NR	*Not reported*	Low
	Kuhlmann 1999	Sleep problems	Valerian	Sleep quality	23 days	short form B	NR	102	NR	NR	NR	NR	*Not reported*	Some concerns
	Jacobs 2005	Insomnia	Valerian	Sleep quality	4 weeks	VAS	higher is better	270	NR	NR	SMD -0.12 (-0.36, 0.12)	NR	*No difference*	Low
	Donath 2000	Insomnia	Valerian	Sleep quality	43 days	response rate	higher is better	16	NR	NR	NR	RR	*Not reported*	Low
	Coxeter 2003	Insomnia	Valerian	Sleep quality	52 days	short form B	NR	21	NR	NR	NR	NR	*Not reported*	Low
	Oxman 2007	Insomnia	Valerian	Sleep quality	4 weeks	VAS (1 to 7)*	higher is better	434	NR	NR	SMD -0.13 (-0.32, 0.07)	NR	*No difference*	Low
	Taibi 2008	Insomnia	Valerian	Sleep quality	44 days	VAS (1 to 9)*	higher is better	16	NR	NR	SMD -0.3 (-1.0, 0.4)	NR	*No difference*	Low
	Footnotes:	* retrieved from primary study												
WHM vs control (no treatment, waitlist, standard care)														
No RCTs reporting this comparison were identified by the eligible reviews														
WHM vs active control														
Shinjyo 2020	Ziegler 2002	Insomnia	Valerian vs Oxazepam	Patient reported improvement	6 weeks	Sleep questionnaire	NR	202	NR	NR	NR	NR	*No difference*	High
	Morin 2005	Insomnia	Valerian + Hops vs Diphenhydramine	HRQoL	4 weeks	NR	NR	184	NR	NR	NR	<0.05	*Favours intervention*	Some concerns
				Sleep quality		Sleep diary	NR		NR	NR	NR	NR	*No difference*	Some concerns
	Maroo 2013	Insomnia	Valerian + Passiflower + Hops) vs Zolpidem	Insomnia severity	2 weeks	ISI	higher is worse	78	NR	NR	NR	<0.05	*Favours intervention*	Some concerns
	Footnotes:													

STUDY RESULTS (as reported by review authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Leach 2015	Ziegler 2002	Insomnia	Valerian vs Oxazepam	Insomnia severity	6 weeks	ISI*	higher is worse	186	NR	NR	SMD 0.13 (-0.16, 0.42)	0.37	*No difference*	High
	Footnotes:	*Reported as sleep quality outcome by Leach 2015												

Abbreviations: BDI, Beck Depression Inventory; C, Comparator; FSS, Fatigue severity scale; I, intervention; ISI, insomnia severity scale; NR, not reported; STAI, state trait anxiety index

STUDY RESULTS (as reported by the review authors)

Review ID	RCT	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
WHM vs placebo														
Bach 2016	Kim 2013	idiopathic chronic fatigue	Panax ginseng (1000 or 2000 mg per day)	Fatigue	end of treatment (4 wks)	NRS (0-100)	Higher means worse fatigue	58/29	41.8 (13.2)	48.8 (7.3)	NR	NR	*Favours intervention*	Low
	Etemadifar 2013	multiple sclerosis patients	Panax ginseng	Fatigue	end of treatment (12 wks)	Modified FIS (0-36)	Higher means greater impact on daily activities	26/26	23.65 (12.8)	23.69 (12.94)	NR	NR	*Favours intervention*	Low
	Footnotes:	Not clear if results for intervention arm were for 2,000 mg/day of 1,000 mg/day. Presumed that results were based on cummulative mean for both.												
Kim 2020	Hartz 20004	idiopathic chronic fatigue	Siberian ginseng 2000 mg vs placebo	Vitality	end of treatment (8 wks)	RVI (0-100)*	Higher means more energy	26/20	−11.8 (3.93)	−10.2 (3.93)	SMD −0.40 (−0.99, 0.19)	NR	*no difference*	Not assessed
	Footnotes:	*presumed the direction of the scale has been reversed to fit with measures of fatigue.												
	Hartz 2004	idiopathic chronic fatigue	Panax ginseng (800 mg bid po)	Fatigue severity	end of treatment (8 wks)	NR	NR	36/40	NR	NR	NR	<0.05	*Favours intervention*	Some concerns
				Vitality	end of treatment (8 wks)	RVI (0-100)*	Higher means more energy	36/40	NR	NR	NR	<0.05	*Favours intervention*	Some concerns
				Mood/ Anxiety	end of treatment (8 wks)	MASQ	NR	36/40	NR	NR	NR	<0.05	*Favours intervention*	Some concerns
	Kim 2016	idiopathic chronic fatigue	Panax ginseng (50 mg bid po)	Fatigue	end of treatment (4 wks)	CIS	NR	72/77	NR	NR	NR	>0.05	*no difference*	Some concerns

STUDY RESULTS (as reported by the review authors)

Review ID	RCT	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Jin 2020	Hyeong-Geug 2013	idiopathic chronic fatigue	Panax ginseng (1or 2 g qd po)	NR*	end of treatment (4 wks)	NRS	NR	90/30	NR	NR	NR	>0.05	*no difference*	Some concerns
				NR*	end of treatment (4 wks)	VAS	NR	90/30	NR	NR	NR	<0.05	*Favours intervention*	Some concerns
	Lee 2016	idiopathic chronic fatigue	EMGE (500 mg bid po)	Fatigue	end of treatment (4 wks)	VAS	Higher means more fatigue	26/26	NR	NR	NR	<0.01	*Favours intervention*	Low
				Fatigue	end of treatment (4 wks)	RPFS	Higher means more fatigue	26/26	NR	NR	NR	>0.05	*no difference*	Low
				HRQoL	end of treatment (4 wks)	SF-36	Higher means better HRQoL	26/26	NR	NR	NR	>0.05	*no difference*	Low
	Gal 1996	idiopathic chronic fatigue	Panax ginseng (1capsule bid po)	Fatigue score	end of treatment (6 wks)	NR	NR	109/109	NR	NR	NR	0.019	*Favours intervention*	Some concerns
	Footnotes:	*not clear if the measure is for fatigue, pain or other outcome												
WHM vs control (no intervention, waitlist, usual care)														
	No RCTs reporting this comparison were identified by the eligible reviews													
WHM vs other intervention														
	No RCTs reporting this comparison were identified by the eligible reviews													

Abbreviations: bid, *bis in die* (twice daily); C, Comparator; CIS, checklist individual strength; EMGE, enzyme-modified ginseng extract; I, intervention; FIS, fatigue impact scale; MASQ, mood and axiety symtpoms questionaire; mg, milligrams; NR, not reported; NRS, numeric rating scale; po, per oral; RPFS, revised Piper fatigue scale; qd, *quaque die* (once daily); RVI, Rand vitalilty index; VAS, visual analogue scale

STUDY RESULTS (as reported by the study authors)

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
WHM vs placebo														
Kim 2021	Lu 2016	Acne vulgaris	Green tea (oral)	Disease severity	4 weeks	Inflammatory lesion count (higher is worse)		64 (33/31)	-3.4 (3.2)	-2 (0.3)	MD -1.40 (-2.5, -0.3)	0.01	*Favours intervention*	Low
						Non-inflammatory lesion count (higher is worse)			0.00 (0.40)	-0.2 (0.40)	MD 0.2 (0.0, 0.40)	0.05	*Not reported*	Low
	Sharquie 2006	Acne vulgaris	Green tea (2% tea lotion	Disease severity	8 weeks	Inflammatory lesion count (higher is worse)		47 (25/24)	-15.7 (4.9)	-2.5 (2.7)	MD -13.20 (-15.40, -11)	NR	*Favours intervention*	Some concerns
	Yoon 2013 (group a)*	Acne vulgaris	Green tea (1% EGCG)	Disease severity	8 weeks	revised Leeds score (mean change from		34 (17/17)	-3.3 (0.1)	-1.2 (0.2)	MD -2.1 (-2.21, -1.99)	<0.001	*Favours intervention*	Low
				Disease severity		Inflammatory lesion count (higher is worse)			-8.9 (2.6)	-1.00 (1.10)	MD-7.90 (-9.24, -6.56)	<0.001	*Favours intervention*	Low
				Disease severity		Non-inflammatory lesion count (higher is worse)			-38.2 (13.60)	1.90 (4.80)	MD -36.30 (-43.16, -29.44)	<0.001	*Favours intervention*	Low
				Global improvement		VAS (0-10) (mean change from baseline)			-6.5 (1.1)	-1.2 (1.4)	MD -5.30 (-6.15, -4.45)	<0.001	*Favours intervention*	Low
	Yoon 2013 (group b)*	Acne vulgaris	Green tea (5% EGCG)	Disease severity	8 weeks	revised Leeds score (mean change from		35 (18/18)	-3.5 (0.7)	-1.2 (0.2)	MD -2.30 (-2.65, -1.95)	<0.001	*Favours intervention*	Low
				Disease severity		Inflammatory lesion count (higher is worse)			-8.6 (1)	-1.00 (1.10)	MD -7.6 (-8.31, -6.89)	<0.001	*Favours intervention*	Low
				Disease severity		Non-inflammatory lesion count (higher is worse)			-31.20 (10)	-1.90 (4.80)	MD -29.30 (-34.57, -24.03)	<0.001	*Favours intervention*	Low
				Global improvement		VAS (0-10) (mean change from baseline)			-5.1 (1.3)	-1.2 (1.4)	MD -3.90 (-4.81, -2.99)	<0.001	*Favours intervention*	Low
	Footnotes:	*Yoon 2013 examined two different concentrations of green tea extract (1% EGCG and 5% EGCG) independently in a split-face trial design.												

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Vaughn 2016	Lalla 2001*	Acne vulgaris	Curcumin combination (oral+topical cream)	Disease severity	4 weeks	Leeds technique (4-point scale)		53 (NR)	NR	NR	NR	NR	*Favours intervention*	High
			Curcumin combination (oral+topical gel)	Disease severity		Leeds technique (4-point scale)			NR	NR	NR	NR	*Favours intervention*	High
	Footnotes:	*The study appears to have several groups in varying combination of oral + topical gel, oral + topical cream, oral + placebo. There is only 1 participant in the placebo + placebo group.												
WHM vs inactive control														
Tuong 2015	Jung 2012	Acne vulgaris	Green tea extract (20mg/mL) vs no intervention	Disease severity	8 weeks	Total acne lesion count (higher is worse)		30 (NR)	NR	NR	NR	p<0.05	*Favours intervention*	High
	Footnotes:													
WHM vs active control														
Kim 2021	Waranuch 2019	Acne vulgaris	Hydrogel* vs 1% clindamycin	Disease severity	4 weeks	Acne severity index		60 (30/30)	-2.00 (4.93)	-1.8 (7.12)	MD -1.30 (-4.9, 2.3)	0.48	*No difference*	Low
						Inflammatory lesion count (higher is worse)			-2.00 (4.93)	1.2 (4.38)	MD -0.8 (-3.16, 1.56)	0.51	*No difference*	Low
	Footnotes:													
	Sharquie 2008	Acne vulgaris	Green tea (2% tea lotion) vs 5% zinc sulfate solution	Disease severity	8 weeks	Inflammatory lesion count (higher is worse)		40 (20/20)	-26.30 (3)	-9.3 (1.3)	MD -17 (-18.43, -15.57)	NR	*Favours intervention*	High
		Footnotes:	* consists of aloe, mangosteen, green tea											

Review ID	Study ID	Condition	Intervention	Outcome	Timing	Outcome measure	measure details	# participants (I/C)	(intervention) n/N (%) or mean (SD)	(comparator) n/N (%) or mean (SD)	Point estimate (95% CI)	*p* -value	*direction of effect*	RoB
Enrst 2002	Basset 1990	Acne vulgaris	Tea tree oil vs benzoyl peroxide	Disease severity	3 months	Global improvement in inflammatory lesions (measure NR)		124	NR	NR	NR	p<0.05	*Favours comparator*	High
	Fulton 1990*	Acne vulgaris after dermabrasion	Aloe vera gel vs polyethylene oxide gel	Global improvement	5 days	Wound healing (reepithelialisation)		18	NR	NR	NR	NR	*Favours intervention*	High
	Footnotes:	*the study is a split-face trial design.												
Vogler 1999	Fulton 1990*	Acne vulgaris after dermabrasion	Aloe vera gel vs polyethylene oxide gel	Global improvement	5 days	Wound healing (reepithelialisation)		18	90% complete	40% to 50% complete	NR	NR	*Favours intervention*	High
	Footnotes:	*the study is a split-face trial design.												

Abbreviations: C, Comparator; I, intervention; IBS, irritable bowel syndrome; NR, not reported