

# Efficacy of journaling in the management of mental illness: a systematic review and meta-analysis

Monika Sohal <sup>1</sup>, Pavneet Singh,<sup>1</sup> Bhupinder Singh Dhillon,<sup>1</sup> Harbir Singh Gill <sup>1,2</sup>

**To cite:** Sohal M, Singh P, Dhillon BS, *et al.* Efficacy of journaling in the management of mental illness: a systematic review and meta-analysis. *Fam Med Com Health* 2022;**10**:e001154. doi:10.1136/fmch-2021-001154

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/fmch-2021-001154>).

## ABSTRACT

**Objectives** Journaling is a common non-pharmacological tool in the management of mental illness, however, no clear evidence-based guideline exists informing primary care providers on its use. We seek here to present this synthesis that may begin to inform future research and eventual evidence-based guideline development.

**Design** Of the 3797 articles retrieved from MEDLINE, EMBASE, PsycINFO, 20 peer-reviewed randomised control trials (31 outcomes) met inclusion criteria. These studies addressed the impact of a journaling intervention on PTSD, other anxiety disorders, depression or a combination of the aforementioned.

**Eligibility criteria** Peer reviewed, randomised control trials on the impact of journaling on mental illness were included.

**Information sources** MEDLINE, EMBASE and PsycINFO.

**Results** The data are highly heterogeneous (control arm= $I^2$  of 71.2%, intervention arm= $I^2$  of 83.8%) combined with a B-level Strength of Recommendation Taxonomy recommendation. It was additionally found that there is a significant pre–post psychometric scale difference between control (–0.01, 95% CI –0.03 to 0.00) and intervention arms (–0.06, 95% CI –0.09 to –0.03). This 5% difference between groups indicates that a journaling intervention resulted in a greater reduction in scores on patient health measures. Cohen's *d* effect size analysis of studies suggests a small to moderate benefit.

**Conclusion** Further studies are needed to better define the outcomes. Our review suggests that while there is some randomised control data to support the benefit of journaling, high degrees of heterogeneity and methodological flaws limit our ability to definitively draw conclusions about the benefit and effect size of journaling in a wide array of mental illnesses. Given the low risk of adverse effects, low resource requirement and emphasis on self-efficacy, primary care providers should consider this as an adjunct therapy to complement current evidence-based management.

## INTRODUCTION

Although journaling is a widely used tool in the management of mental illness, there have been limited efforts to systematically review this paradigm in the Canadian context. There is additionally no current evidence-based guidance about the utility of

journaling as a non-pharmacological treatment modality for family physicians. This is despite the fact that general practitioners represent a large majority of the front line of treatment for individuals suffering from mental illness.<sup>1</sup>

Similar to anything else that is prescribed, journaling as an intervention has several complexities which are important to consider when using this tool.<sup>2</sup> These complexities include variables such as what is written, instruction provided on how to journal, duration of a journaling session and the optimal number of journaling sessions needed to see improvement.<sup>3</sup> In respect to medications, the duration of effect with respect to pharmacokinetics is well understood among the scientific community.<sup>4</sup> However, there is little known and understood about the duration of effect for journaling.

The two primary forms of journaling used in psychotherapy are expressive writing and gratitude journaling.<sup>5</sup> Expressive writing is a journaling technique which is performed for 3–4 sessions about 'one's deepest thoughts and feelings' for 20 min per session.<sup>6</sup> A gratitude journal on the other hand is a diary of accounts for which one is grateful, and where attention is focused on the positive aspects of one's life.<sup>7</sup> Instruction on how to journal is essential in ensuring proper utility of the tool and instilling self-efficacy in the patient.<sup>3</sup> The literature, however, is sparse in areas investigating these variables which moderate the effects of journaling.

This is the first study to examine the efficacy of journaling as an intervention for improving mental health outcomes with a meta-analysis.

We seek to state in terms relevant to primary care clinicians the current state of the data and provide further areas for study. To the best of our knowledge, this paper is the first to report a systematic review and



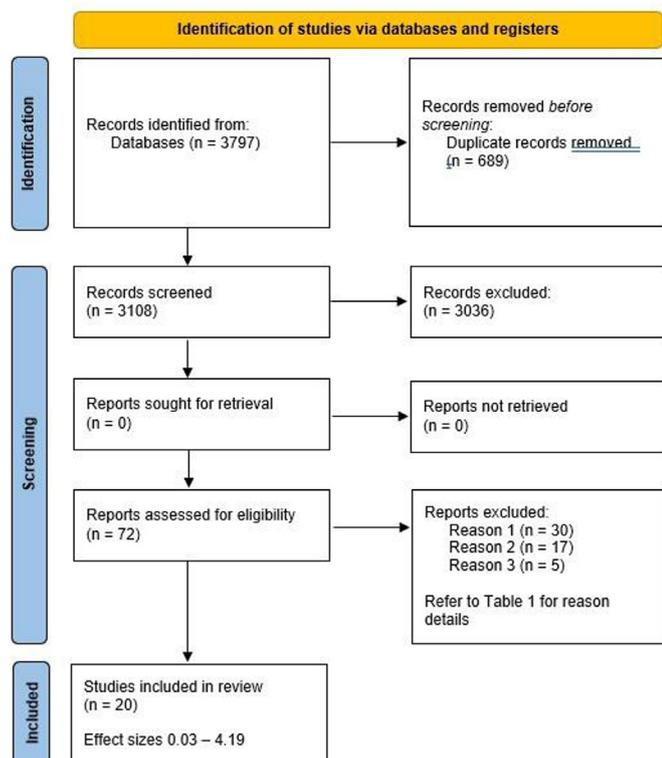
© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Northeast Addiction and Mental Health Centre for Holistic Recovery, Calgary, Alberta, Canada

<sup>2</sup>Department of Family Medicine, University of Calgary, Calgary, Alberta, Canada

### Correspondence to

Dr Harbir Singh Gill;  
hsgill@namrecovery.com



**Figure 1** Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 flow diagram.

meta-analysis of journaling as an intervention to improve mental health.

## METHODS

### Study design and search strategy

A systematic review and meta-analysis of the literature focused on journaling interventions to improve mental health outcomes was conducted. The different phases of the systematic review are displayed in [figure 1](#).

The following databases were searched: MEDLINE, EMBASE and PsycINFO (March 2020). The main search concepts used as both subject headings and keywords journaling and mental health. Detailed MEDLINE search strategy can be seen in online supplemental appendix 1. Initial screening of titles and abstracts was done by MS and BSD. The initial screening involved reading titles and abstracts.

Articles were included and excluded based on the criteria outlined in [table 1](#). Secondary screening involved reading full-length articles and was done by MS, BSD and PS. Disagreements over study eligibility were resolved through discussion with HSG. The list of selected articles can be seen in online supplemental appendix 2.

## STATISTICAL ANALYSIS

### Meta-analysis statistics

The effect measure was calculated by subtracting post-study scores from baseline scores and then dividing this number by the total maximum score allocatable on

**Table 1** Inclusion and exclusion criteria used in primary and secondary screenings

Inclusion criteria	Exclusion criteria
Peer reviewed	Studies which examined the effect of multiple types of journaling
Randomised control trials	Studies with insufficient or missing data which was required to perform a meta-analysis
Studied the impact of a journaling intervention on mental illness	Studies which examined the effects of a journaling intervention on well-being only

the questionnaire used, to assess percentage change. This method was selected, as the majority of studies in this review implemented different measures for scoring changes in mental health status. All questionnaires and health measures used in the studies included in this review indicate that a higher score is suggestive of more severe mental illness. MEDCALC Software was used to calculate 95% CIs and SE for the effect measures. A random effects linear meta-regression analysis was performed on both the intervention and control groups. A forest plot was generated using STATA V.13 software.

### Meta-regression statistics

Meta-regression, a form of moderator analysis, was used to determine sources of heterogeneity within the sample of studies and to test whether the relationship between two variables depends on (is moderated by) the value of a third variable. The analysis was performed on the following variables: type of journaling, intervention duration, journal collection/analysis, sample size, participant sex and age, and geographical region where the studies were conducted. Additionally, type of mental illness was also investigated as a source of heterogeneity with anxiety as the reference category in relation to PTSD and depression. The meta-regression was conducted on all included articles, as well as the symptom subgroups of anxiety, PTSD and depression individually.

### Cohen's d effect size

Our preferred manner of stating the results from this review was in an ORs and number needed to treat. However, the data did not present in a manner that allowed for us to calculate these numbers. Given our inability to calculate a proper OR for the studies, we calculated a Cohen's d to give a perspective of significance. Cohen's d effect size was calculated using means, SD, and sample sizes of the intervention group at baseline and post-intervention. Three particular reference points have been established for interpreting Cohen's d effect size. Effect sizes are classified as small ( $d=0.2$ ), medium ( $d=0.5$ ) and large ( $d\geq 0.8$ ).<sup>8</sup> According to Cohen, 'a medium effect of 0.5 is visible to the naked eye of a careful observer'.<sup>8</sup> A small effect of 0.2 is noticeably smaller than medium but

not so small as to be trivial. A large effect of 0.8 is the same distance above the medium as small is below it. Increasing effect sizes indicate a larger difference between the means of the intervention and control groups.

### Risk of bias assessment

The revised Cochrane risk of bias tool (ROB-2)<sup>9</sup> for randomised control trials was used to assess outcomes at the study level for each article included in this review. This tool scores studies on five domains including randomisation (domain 1), effect of assignment to the intervention (domain 2), missing outcome data (domain 3), measurement of the outcome (domain 4), selection of the reported result (domain 5), and lastly, the overall risk of bias.<sup>9</sup> Please see Sterne *et al*'s paper<sup>9</sup> for more information on the ROB-2 tool.

### Strength of Recommendation Taxonomy

The Strength of Recommendation Taxonomy (SORT) tool was used to determine the strength and quality of evidence for the studies included in this systematic review and meta-analysis.

The SORT has three levels of recommendation of evidence including A-level, B-level and C-level. The definitions of each SORT level of recommendation are quoted below as per Ebell *et al*.<sup>10</sup>

'An A-level recommendation is based on consistent and good quality patient-oriented evidence. A B-level recommendation is based on inconsistent or limited quality patient-oriented evidence. A C-level recommendation is based on consensus, usual practice, opinion, disease-oriented evidence, or case series for studies of diagnosis, treatment, prevention or screening.'<sup>10</sup>

## RESULTS

### Search outcomes

An initial search of the databases generated a total of 3797 articles. A total of 689 duplicate articles were removed from the initial pool of articles and 3108 articles remained for primary screening. The primary and secondary screenings resulted in a total of 20 articles which were analysed for the purposes of this meta-analysis. Online supplemental appendices 3,5 for details the demographic variables, results, questionnaires and measure used to calculate outcomes across studies.

A total of 20 articles and 31 outcomes are presented in this paper. There are more outcomes than articles as several articles examined the effects of a journaling intervention on more than one symptom of mental illness (eg, anxiety and depressive symptoms, or anxiety, depression and PTSD). All articles employed either an expressive writing (17) or gratitude journaling (3) intervention. As seen in [table 2](#), most studies used a different measure to determine changes in mental health. In order to standardise each outcome, the postintervention scores were subtracted from baseline scores and then divided by the maximum score possible on the respective questionnaire

to determine percentage change. Data from the intervention and control groups were calculated and analysed separately. Online supplemental appendix 4 details the journaling type, mental illness symptom studied, effect size and significance of outcomes across individual studies.

The systematic review revealed 68% of the intervention outcomes were effective, with a significant difference between the control and intervention groups supporting the efficacy of journaling. As shown in [table 3](#), of the nine outcomes regarding PTSD (all employed an expressive writing intervention), six showed significant reductions in symptoms postintervention. Of the four outcomes that involved a gratitude journaling intervention, three showed significant improvements in symptomology postintervention. Of the 27 outcomes that implemented an expressive writing intervention, 19 showed significant improvements postintervention.

As seen in [table 2](#), a considerable majority (85%) of the studies implemented a short-term journaling intervention ranging from 2 to 4 journaling sessions. Three studies implemented a long-term journaling intervention ranging from 7 to 32 journaling sessions. Additionally, 55% of the studies collected and/or analysed participant journal entries. Out of the 20 studies, none implemented a preparatory/educational session at the preintervention phase.

[Table 3](#) summarises the results from the meta-regression analyses performed on multiple study variables, examining also the three symptom subgroups of anxiety, depression and PTSD. Variables assessed to determine moderators of heterogeneity included: type of journaling, length of intervention, age, region, sex, mental illness symptom (depression and PTSD with anxiety as the reference category), sample size (<30 or >30 participants), and journals analysis and/or collection. Of these variables, the only significant moderator was journal analysis and/or collection ( $r=-0.064$ ;  $p=0.021$ ), which indicates that when journals are not collected or analysed the efficacy of journaling is greater. The remaining variables were not deemed as significant moderators contributing to the heterogeneity of the data for the whole group.

We explored this further in a meta-regression on symptom subgroups, please refer to tabulated summaries in online supplemental appendix 6. We found that the anxiety subgroup showed poorer pre/post scores for both intervention and control arms for studies with mixed genders compared with a single gender (intervention  $r=0.178$   $p=0.038$ , control  $r=0.181$   $p=0.002$ ). From this, we can infer that journaling works better for women and this improvement in health outcomes is also supported in the literature.<sup>6</sup>

Depression subgroup analysis showed that journaling length >30 days compared with <30 days improved pre-post scores by 10.4% ( $r=-0.104$ ,  $p=0.005$ ), indicating a longer intervention is optimal.

PTSD subgroup showed age as a relevant moderator for the intervention arm ( $r=0.073$ ,  $p=0.012$ ), whereby pre/

**Table 2** Intervention parameters including journaling duration, journal analysis and/or collection and preparatory session administration preintervention

Study	Journaling session duration (min)	No of journaling sessions	Journals collected and/or analysed (yes/no)	Preparatory session* (yes/no)
Alparone <i>et al</i> (2015)	20	3	Yes	No
Bernard <i>et al</i> (2006)	15	3	Yes	No
Dennick <i>et al</i> (2015)	20	3	Yes	No
Di Blasio <i>et al</i> (2015)	Unspecified	2	No	No
Ducasse <i>et al</i> (2018)	Unspecified	7	No	No
Graf <i>et al.</i> (2008)	20	2	No	No
Horsch <i>et al</i> (2016)	15	3	Yes	No
Jensen-Johansen <i>et al</i> (2012)	20	3	No	No
Barton & Jackson (2008)	20	3	Yes	No
Koopman <i>et al</i> (2005)	20	4	No	No
Barry and Singer (2001)	20	4	Yes	No
Lovell <i>et al</i> (2016)	20	3	Yes	No
Suhr, Risch and Wilz (2017)	Unspecified	3	No	No
Martino <i>et al</i> (2012)	Unspecified	2	No	No
Meshberg-Cohen <i>et al</i> (2014)	20	4	Yes	No
Possemato <i>et al</i> (2010)	15	3	Yes	No
Rabiepoor <i>et al</i> (2019)	15	21	No	No
Rawlings <i>et al</i> (2018)	20	4	Yes	No
Schache <i>et al</i> (2019)	Unspecified	32	No	No
Wong and Mak (2016)	20	3	Yes	No

\*A preparatory session refers to an educational session given to participants preintervention which outlines the mechanisms and theories underlying the efficacy of journaling as well as how to journal in the most effective manner.

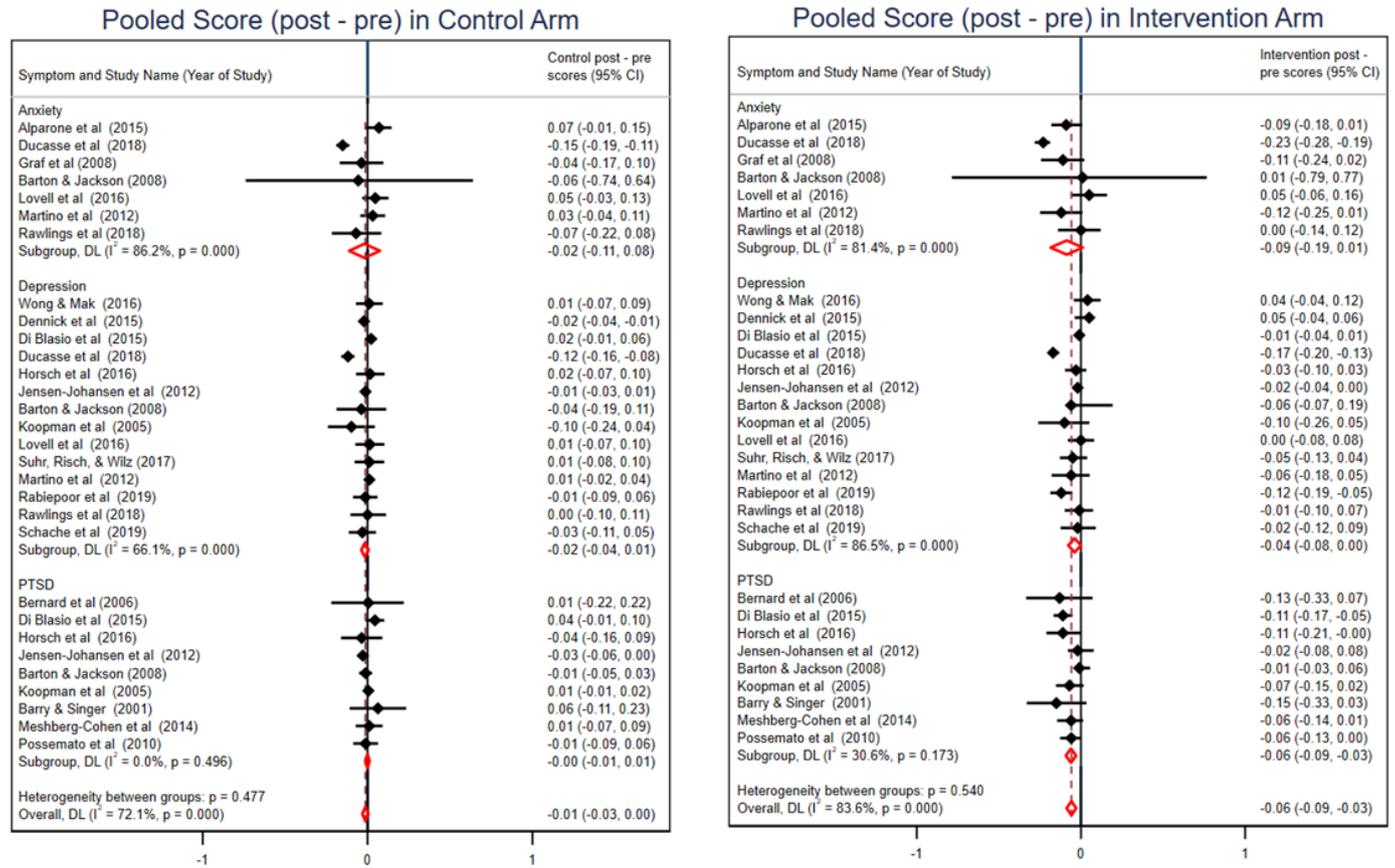
**Table 3** Results of meta-regression without subgroups

Study arm	Intervention arm		Control arm		
	Regression coefficient	P value	Regression coefficient	P value	
Sample size	0.023	0.445	-0.025	0.247	
Type of journaling	-0.043	0.176	-0.036	0.124	
Study duration	-0.052	0.183	-0.043	0.147	
Journal analysis	-0.064	0.021*	-0.026	0.254	
Sex	0.052	0.063	0.030	0.158	
Age	0.019	0.503	-0.034	0.111	
Region-Europe	-0.097	0.217	-0.023	0.714	0.971†
Region-North America	-0.122	0.150	-0.022	0.750	
Region-Australia	-0.060	0.596	-0.042	0.635	
Depression‡	0.060	0.122	0.006	0.838	0.763
PTSD‡	0.030	0.459	0.022	0.51	

\*Sample size <30 or >30.

†When regions analysed in aggregate.

‡Depression and PTSD regression analyses were calculated using anxiety as the reference category. PTSD, post-traumatic stress disorder.



**Figure 2** Standardised forest plot for control and intervention groups.

post scores worsened with age: journaling is less effective the older one gets.

The forest plot in figure 2 illustrates that following standardisation of the data there is a significant difference between control ( $-0.01$ , 95% CI  $-0.03$  to  $0.00$ ,  $p \leq 0.001$ ) and intervention arms ( $-0.06$ , 95% CI  $-0.09$  to  $-0.03$ ,  $p \leq 0.001$ ) between all groups. This 5% difference between groups is statistically significant and indicates that a journaling intervention resulted in a greater reduction in scores on patient health measures when compared with the control arm.

Subgroup analysis further indicated that when studies examining journaling in anxiety disorders, a reduction in pre-post intervention scores was 9%–0.09 (95% CI  $-0.19$  to  $0.01$ ,  $p \leq 0.001$ ) compared with control scores at 2%–0.02, (95% CI  $-0.11$  to  $0.08$ ,  $p \leq 0.001$ ).

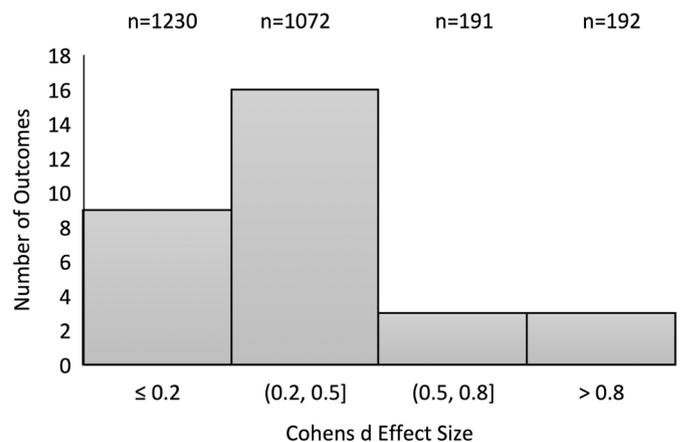
In the depression subgroup, the intervention scores difference was 4%–0.04 (95% CI  $-0.08$  to  $0.00$ ,  $p \leq 0.001$ ) compared with the control scores 2%–0.02 (95% CI  $-0.04$  to  $0.01$ ,  $p \leq 0.001$ ).

In the PTSD subgroup, intervention scores showed a difference of 6%–0.06 (95% CI  $-0.09$  to  $-0.03$ ,  $p \leq 0.001$ ) compared with control scores (95% CI  $-0.01$  to  $0.01$ ,  $p = 0.496$ ).

The data, however, have a high degree of heterogeneity indicated by an  $I^2$  of 71.2% in the control arm and an  $I^2$  of 83.8% in the intervention arm. There was lower heterogeneity in the PTSD subgroup at  $I^2$  of 30.6%,

in comparison to anxiety ( $I^2 = 86.5\%$ ) and depression ( $I^2 = 81.4\%$ ) subgroups.

Figure 3 illustrates a histogram depicting Cohen's d effect size<sup>8</sup> for the 31 outcomes. Cohen's d effect size provides a standardised representation of the distribution of means between the control and intervention group. For the purpose of calculating Cohen's d effect size for this histogram, postintervention scores of the control and intervention groups were used. A total of nine outcomes indicated a Cohen's effect size  $< 0.2$  (small effect size).



**Figure 3** Histogram of Cohen's d effect size of journaling efficacy. n indicates the numbers of participants within each Cohen's effect size classification.

Sixteen outcomes fall between 0.2 and 0.5 (small to moderate effect size) and the remaining six outcomes are above 0.5 (moderate to large effect size).

The ROB-2 tool<sup>9</sup> was used to determine the risk of bias at the individual study level for all articles included in this review. The majority of the studies (12) fall under a 'low-risk' categorisation with three falling under a 'high-risk' categorisation and five falling under the 'some concerns' categorisation. Refer to online supplemental appendix 7 for a detailed presentation of ROB-2 results.

The SORT taxonomy was used to determine quality of evidence of all studies included in this article. The articles included in this review are graded as a B-level recommendation. A B-level recommendation indicates that 'the data are inconsistent, or limited quality patient-oriented evidence'.<sup>10</sup>

## DISCUSSION

This is the first study to examine the efficacy of journaling as an intervention for improving mental health outcomes with a meta-analysis. The results of this work revealed that a journaling intervention resulted in an average statistically significant 5% reduction in patient scores on mental health measures compared with control arms, with a greater benefit in anxiety (9%) and PTSD (6%) symptom subgroups, and a lesser benefit in depression subgroup (2%). Due to heterogeneity of the scales used, we are unable to comment on whether this finding is clinically significant (as that varies depending on the scale used), though our findings would support the wide use of journaling in many forms of psychotherapy for a broad range of mental illnesses.<sup>11</sup>

The high degree of heterogeneity of the data in the intervention arm ( $I^2=83.8\%$ ) as well as the B-level SORT recommendation indicates that it is difficult to draw any robust conclusions from this meta-analysis, although the PTSD subgroup had moderate heterogeneity ( $I^2=30.6\%$ ) which allows for greater confidence in the 6% reduction of symptom scales seen with journaling for these patients.

The meta-regression yielded some clues as to sources of heterogeneity including study population, age and gender composition, length of time of journaling intervention (<30 days or >30 days), and whether journals were collected/analysed. The identification of these variables as sources of heterogeneity in this meta-analysis demonstrates the need for a mixed gender/age population that represents family physician clinical populations in future studies, as well as encouraging journaling intervention >30 days in duration and recommending against collecting journals.

This last factor of journal collection and/or analysis is critical to acknowledge given this was revealed to be a statistically significant moderator. Eleven studies (18 outcomes) collected and/or analysed participant journals. This would likely have adversely affected results. Much of the efficacy of journaling comes from the ability of one to write openly and honestly, without the fear of

judgement, safe in the knowledge that the journal will not be read by others.<sup>12</sup> While some involved in those studies were only collecting journals to ensure compliance, some were in fact analysed for content. Both practices in our opinion are counterproductive to optimal journaling practices.

We also note that a detailed preparatory/educational training session is vital in allowing the patient to understand why journaling works and how to effectively journal.<sup>3</sup> Although all 20 studies in this paper provided some limited written and/or verbal guidance, no study implemented a training session or provided participants with an opportunity to practice with a trained instructor or ask questions. It could be argued that any efficacy that may be attributable to journaling may well be diminished due to the lack of a training/preparatory session. Journaling is an intervention that requires education, coaching and practice, in order to be undertaken in a manner that ensures optimal effects.<sup>2</sup>

We intended to calculate an OR or 'number needed to treat' for journaling, however, the data did not present in a manner that allowed us to do so. Due to differences in study designs, comparisons of behaviour change/non-pharmacological interventions (eg, journaling) to conventional pharmacological treatments such as antidepressant efficacy are challenging. However, we were able to find a meta-analysis and systematic review done by Fournier *et al*<sup>13</sup> on antidepressant efficacy using Cohen's *d* effect size. Fournier's findings depict that most antidepressant therapy studies show a moderate effect size and antidepressants are most effective in a population suffering from very severe depression. Interestingly, our findings for journaling show small to moderate effect sizes, suggesting that journaling interventions show promise to become a highly efficacious evidence-based therapy once more well-designed trials are completed.

This meta-analysis provides further affirmation that journaling as an intervention has merit and can be an efficacious adjunct when prescribed and implemented properly. The findings of this study can be applied in primary care practice by using journaling as a low risk, low-resource intensive adjunct to standard therapy for patients with mental health concerns. This study can also guide future research by highlighting the criteria required for robust studies. Journaling is a promising intervention for the treatment of mental health disorders but requires more robust high-quality longitudinal studies to appropriately assess its efficacy. In addition to this, participants require adequate preparation and the opportunity to practice if we are to fully investigate how journaling could benefit patients.<sup>2</sup> We also recommend not sharing the participants' personal journal with the study team.

Even though the included studies were found to have a low risk of bias, one of the limitations of this meta-analysis include a lack of high-quality studies with low heterogeneity, thus curbing the ability to draw clinically relevant conclusions. In addition, we only included randomized controlled trials (RCTs) in this study and

thus, may have missed important studies on journaling which used different study designs. Finally, because meta-regression was performed within subgroups (anxiety, depression and post-traumatic stress disorder (PTSD)), the number of studies within subgroups became less than 10 across several groups, and meta-regression is generally not recommended with fewer than ten studies in a meta-analysis.

## CONCLUSION

Journaling is an adjunct low-cost, low-side effect therapy that can help family physicians in the management of common mental health symptoms that is supported by randomised controlled trials as summarised in this systematic review with meta-analysis. While study quality is overall low, the cost–benefit ratio is in favour of family physicians advocating for use of this modality as an adjunct to other therapies for common mental health conditions. Further studies are required to better evaluate the parameters that can further optimise journaling efficacy as an intervention on mental health outcomes.

**Acknowledgements** Mohammad Ziaul Islam Chowdhury (PhD Candidate) - tasks performed: statistical analysis; Gursharan Kaur Dalam - tasks performed: data extraction; Melissa Marshall, PhD (final manuscript revision); Chantal Hansen, MGIS (final manuscript revision).

**Contributors** MS: planning, data collection, data analysis, synthesis, editing. PS: planning, data collection. BSD: planning, data analysis, synthesis. HG: concept, editing, final revision.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability

of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iDs

Monika Sohal <http://orcid.org/0000-0003-4485-748X>

Harbir Singh Gill <http://orcid.org/0000-0003-1554-1064>

## REFERENCES

- 1 Wang J, Langille DB, Patten SB. Mental health services received by depressed persons who visited general practitioners and family doctors. *Psychiatr Serv* 2003;54:878–83.
- 2 Hayman B, Wilkes L, Jackson D. Journaling: identification of challenges and reflection on strategies. *Nurse Res* 2012;19:27–31.
- 3 Hubbs DL, Brand CF. The paper mirror: understanding reflective Journaling. *J Exp Educ* 2005;28:60–71.
- 4 Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* 2018;391:1357–66.
- 5 Wong YJ, Owen J, Gabana NT, et al. Does gratitude writing improve the mental health of psychotherapy clients? Evidence from a randomized controlled trial. *Psychother Res* 2018;28:192–202.
- 6 Pennebaker JW, Chung CK. Expressive writing and its links to mental and physical health. *Oxford Handb Heal Psychol* 2011.
- 7 Waters L. A review of school-based positive psychology interventions. *Aust Educ Dev Psychol* 2011.
- 8 Lerner AJ. Effect size (Cohen's D) of cognitive screening instruments examined in pragmatic diagnostic accuracy studies. *Dement Geriatr Cogn Dis Extra* 2014;4:236–41.
- 9 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:14898.
- 10 Ebell MH, Siwek J, Weiss BD. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. *J Am Board Fam Pract* 2004;69:548–56.
- 11 Ullrich PM, Lutgendorf SK. Journaling about stressful events: effects of cognitive processing and emotional expression. *Ann Behav Med* 2002;24:244–50.
- 12 Phipps JJ. E-journaling: achieving interactive education online. *Educ Q* 2005.
- 13 Fournier JC, DeRubeis RJ, Hollon SD, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010;303:47–53.

## APPENDIX I

### Detailed search strategy and selection methods

1. All databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.

MEDLINE: date accessed – March 5<sup>th</sup> 2020

EMBASE: date accessed – March 5<sup>th</sup> 2020

PsycInfo: date accessed – March 5<sup>th</sup> 2020

2. Search strategy for MEDLINE (the same strategy was used for PsycInfo and EMBASE)

	Search terms	# of articles retrieved
1	journaling.mp.	368
2	gratitude.mp.	1366
3	Reflective journal*.mp.	319
4	Journal writing.mp.	98
5	Reflective practice.mp.	1227
6	Expressive writing.mp.	311
7	Written emotional disclosure.mp.	61
8	exp Depression/	115331
9	depress*.mp.	532094
10	exp Self-Injurious Behavior/	69278
11	exp Substance-Related Disorders/	273882
12	addiction.mp.	45821
13	exp Anxiety/ or exp Anxiety Disorders/	152206
14	exp Mental Health/	36661
15	exp Stress Disorders, Post-Traumatic/	31689
16	exp Mental Disorders/	1216337
17	trauma.mp.	259553
18	exp Chronic Pain/	13647
19	mental health.mp.	184811
20	1 or 2 or 3 or 4 or 5 or 6 or 7	3545
21	8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19	1989928
22	20 and 21	702
23	limit 22 to English language	686
	<b>Total</b>	<b>686</b>
	<b>duplicates</b>	<b>0</b>

3. Methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.

Excel was utilized to extract articles from the databases and for the primary screening process. Inclusion criteria was determined following a brief literature search prior to article retrieval and was modified accordingly with team input as the screening proceeded. Two reviewers performed the primary screen in collaboration with a third reviewer available to resolve any conflict regarding inclusion. Two reviewers performed the secondary screen in collaboration with a third reviewer available to resolve any conflict regarding inclusion

4. Methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.

Two reviewers extracted data from included studies independently and in duplicate, and a third reviewer was available to resolve any conflict regarding inclusion

5. Methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.

The Cochrane risk of bias tool (ROB-2) and strength of recommendations taxonomy (SORT) were utilized to determine risk of bias and quality of the literature, respectively. A single reviewer completed both of these with reference to guidelines established by the ROB-2 and SORT creators.

## APPENDIX II

### References for Articles Analyzed in the Systematic Review and Meta-Analysis

1. Alparone FR, Pagliaro S, Rizzo I. The words to tell their own pain: Linguistic markers of cognitive reappraisal in mediating benefits of expressive writing. *J Soc Clin Psychol*. 2015;34(6):495-507. doi:10.1521/jscp.2015.34.6.495
2. Barry LM, Singer GHS. After the NICU Experience Through Journal Writing. 2015:287-297.
3. Barton K, Jackson C. Impact of Writing About Caregiving Experiences. *Aust N Z J Psychiatry*. 2008;42:693-701.
4. Bernard M, Jackson C, Jones C. Written emotional disclosure following first-episode psychosis: Effects on symptoms of post-traumatic stress disorder. *Br J Clin Psychol*. 2006;45(3):403-415. doi:10.1348/014466505X68933
5. Dennick K, Bridle C, Sturt J. Written emotional disclosure for adults with Type 2 diabetes: a primary care feasibility study. *Prim Health Care Res Dev*. 2015;16(2):179-187. doi:10.1017/S1463423614000188
6. Di Blasio P, Camisasca E, Caravita SCS, Ionio C, Milani L, Valtolina GG. The effects of expressive writing on postpartum depression and posttraumatic stress symptoms. *Psychol Rep*. 2015;117(3):856-882. doi:10.2466/02.13.PR0.117c29z3
7. Ducasse D, Dassa D, Courtet P, et al. Gratitude diary for the management of suicidal inpatients: A randomized controlled trial. *Depress Anxiety*. 2019;36(5):400-411. doi:10.1002/da.22877
8. Horsch A, Tolsa JF, Gilbert L, du Chêne LJ, Müller-Nix C, Graz MB. Improving Maternal Mental Health Following Preterm Birth Using an Expressive Writing Intervention: A Randomized Controlled Trial. *Child Psychiatry Hum Dev*. 2016;47(5):780-791. doi:10.1007/s10578-015-0611-6
9. Jensen-Johansen MB, Christensen S, Valdimarsdottir H, et al. Effects of an expressive writing intervention on cancer-related distress in Danish breast cancer survivors - Results from a nationwide randomized clinical trial. *Psychooncology*. 2013;22(7):1492-1500. doi:10.1002/pon.3193
10. Koopman C, Ismailji T, Holmes D, Classen CC, Palesh O, Wales T. The effects of expressive writing on pain, depression and posttraumatic stress disorder symptoms in survivors of intimate partner violence. *J Health Psychol*. 2005;10(2):211-221. doi:10.1177/1359105305049769
11. Lovell B, Moss M, Wetherell MA. Assessing the feasibility and efficacy of written benefit-finding for caregivers of children with autism: A pilot study. *J Fam Stud*. 2016;22(1):32-42. doi:10.1080/13229400.2015.1020987
12. Martino ML, Freda MF, Camera F. Effects of Guided Written Disclosure Protocol on mood states and psychological symptoms among parents of off-therapy acute lymphoblastic leukemia children. *J Health Psychol*. 2013;18(6):727-736. doi:10.1177/1359105312462434
13. Meshberg-Cohen S, Svikis D, McMahon TJ. Expressive Writing as a Therapeutic Process for Drug-Dependent Women. *Subst Abus*. 2014;35(1):80-88. doi:10.1080/08897077.2013.805181

14. Possemato K, Ouimette P, Geller PA. Internet-based expressive writing for kidney transplant recipients: Effects on posttraumatic stress and quality of life. *Traumatology (Tallahass Fla)*. 2010;16(1):49-54. doi:10.1177/1534765609347545
15. Rabiepoor S, Vatankhah-Alamdary N, Khalkhali HR. The Effect of Expressive Writing on Postpartum Depression and Stress of Mothers with a Preterm Infant in NICU. *J Clin Psychol Med Settings*. 2019;(0123456789). doi:10.1007/s10880-019-09688-2
16. Rawlings GH, Brown I, Stone B, Reuber M. A pilot randomised controlled trial of a home-based writing intervention for individuals with seizures. *Psychol Heal*. 2018;33(9):1151-1171. doi:10.1080/08870446.2018.1478974
17. Schache KR, Hofman PL, Serlachius AS. A pilot randomized controlled trial of a gratitude intervention for adolescents with Type 1 diabetes. *Diabet Med*. 2020;37(8):1352-1356. doi:10.1111/dme.14078
18. Suhr M, Risch AK, Wilz G. Maintaining Mental Health Through Positive Writing: Effects of a Resource Diary on Depression and Emotion Regulation. *J Clin Psychol*. 2017;73(12):1586-1598. doi:10.1002/jclp.22463
19. Wong CCY, Mak WWS. Writing Can Heal: Effects of Self-Compassion Writing among Hong Kong Chinese College Students. *Asian Am J Psychol*. 2016. doi:10.1037/aap0000041

### APPENDIX III

#### Demographic variables, results and questionnaires used to measure outcomes across studies

Study	Population	Sample Size	Results (+/-)*	Measures <sup>±</sup>
Alparone et al (2015)	undergraduate students	62	(+) significant decrease in anxiety	POMS
Barry & Singer (2001)	mothers with an infant in NICU	38	(+) significant decrease in PTSD-related symptoms	IES-R
Barton & Jackson (2008)	caregivers of individuals suffering from psychosis	36	(-) no significant outcomes	GHQ-28, IES-R
Bernard et al (2006)	individuals recovering from psychosis	22	(+) significant decrease in PTSD-related symptoms	IES-R
Dennick et al (2015)	adults with type 2 diabetes	41	(-) significant increase in depressive symptoms	CES-D
Di Blasio et al (2015)	women with PPD	113	(+) significant decrease in symptoms of PPD & PTSD	BDI-II, PPQ
Ducasse et al (2018)	suicidal inpatients	198	(+) significant decrease in symptoms of depression and anxiety	STAI, BDI
Graf et al. (2008)	individuals receiving outpatient psychotherapy	44	(+) significant decrease in symptoms of anxiety	DASS

Horsch et al (2016)	mothers who have had a pre-term birth	54	(+) significant decrease in symptoms of depression and PTSD	EPDS, PPQ
Jensen-Johansen et al (2012)	breast cancer survivors	318	(+) significant decrease in symptoms of depression and PTSD	BDI, IES-R
Koopman et al (2005)	survivors of intimate partner violence	47	(+) significant decrease in depressive symptoms (-) no significant reductions of PTSD-related symptoms	BDI, PCL-5
Lovell et al (2016)	caregivers of children with autism	74	(+) significant decrease in symptoms of depression (-) no significant reductions in symptoms of anxiety	HADS
Martino et al (2012)	parents of children suffering from leukemia	46	(+) significant decrease in symptoms of depression and anxiety	POMS
Meshberg-Cohen et al (2014)	women receiving treatment for addiction	149	(+) significant reductions in symptoms of PTSD	PDS
Possemato et al (2010)	kidney transplant recipients	48	(-) no significant reductions in symptoms of PTSD	PCL-M
Rabiepoor et al (2019)	mothers with infants in the NICU	91	(+) significant decrease in symptoms of PPD	EPDS

Rawlings et al (2018)	individuals suffering from seizures	68	(-) no significant reductions in anxiety (+) significant decrease in depressive symptoms	GAD-7, NDDI-E
Schache et al (2019)	adolescents with type 1 diabetes	60	(-) no significant reductions in symptoms of depression	CES-D
Suhr, Risch & Wilz (2017)	recently discharged psychiatric patients	89	(+) significant decrease in symptoms of depression	BDI-II
Wong & Mak (2016)	university students	65	(-) no significant reductions symptoms of depression	CES-D

\*+ indicates an effective intervention with significant differences between control and intervention groups.

- indicates an intervention that was not effective with no significant differences between control and intervention groups.

± refer to Appendix V for detailed names of measures used.

**APPENDIX IV****Journaling type, mental illness symptom studied, effect size, and significance of outcomes across individual studies**

Study	Intervention	Mental Illness Symptom	Cohen's d Effect Size	Outcome (+/-)*
Alparone et al (2015)	Expressive Writing	Anxiety	0.646	+
Graf et al. (2008)	Expressive Writing	Anxiety	0.293	+
Lovell et al (2016)	Expressive Writing	Anxiety	0.432	-
Martino et al (2012)	Expressive Writing	Anxiety	0.396	+
Rawlings et al (2018)	Expressive Writing	Anxiety	0.143	-
Barton & Jackson (2008)	Expressive Writing	Anxiety	0.026	-
Barry & Singer (2001)	Expressive Writing	Depression	0.805	+
Barton & Jackson (2008)	Expressive Writing	Depression	0.28	-
Dennick et al (2015)	Expressive Writing	Depression	4.193	-
Di Blasio et al (2015)	Expressive Writing	Depression	0.418	+
Horsch et al (2016)	Expressive Writing	Depression	0.221	+

Jensen-Johansen et al (2012)	Expressive Writing	Depression	0.117	+
Koopman et al (2005)	Expressive Writing	Depression	0.376	+
Lovell et al (2016)	Expressive Writing	Depression	0.296	+
Martino et al (2012)	Expressive Writing	Depression		+
Rabiepoor et al (2019)	Expressive Writing	Depression	0.705	+
Rawlings et al (2018)	Expressive Writing	Depression	0.275	+
Wong & Mak (2016)	Expressive Writing	Depression	0.391	-
Barry & Singer (2001)	Expressive Writing	PTSD	0.766	+
Barton & Jackson (2008)	Expressive Writing	PTSD	0.039	-
Bernard et al (2006)	Expressive Writing	PTSD	0.296	+
Di Blasio et al (2015)	Expressive Writing	PTSD	0.937	+
Horsch et al (2016)	Expressive Writing	PTSD	0.371	+
Jensen-Johansen et al (2012)	Expressive Writing	PTSD	0.147	+

Koopman et al (2005)	Expressive Writing	PTSD	0.104	-
Meshberg-Cohen et al (2014)	Expressive Writing	PTSD	0.000	+
Possemato et al (2010)	Expressive Writing	PTSD	0.411	-
Ducasse et al (2018)	Gratitude Journaling	Anxiety	0.279	+
Ducasse et al (2018)	Gratitude Journaling	Depression	0.104	+
Schache et al (2019)	Gratitude Journaling	Depression	0.16	-
Suhr, Risch & Wilz (2017)	Gratitude Journaling	Depression	0.33	+

\*+ indicates an effective intervention with significant differences between control and intervention groups.

- indicates an intervention that was not effective with no significant differences between control and intervention groups.

## APPENDIX V

### Abbreviated and full names of patient health measures used in studies to measure outcomes

<b>Measures</b>
BDI = Beck Depression Inventory
BDI-II = Beck Depression Inventory second edition
CES-D = Centre for Epidemiological Studies Depression Scale
DASS = Depression Anxiety Stress Scale
EPDS = Edinburgh Post-natal Depression Scale
GAD-7 = General Anxiety Disorder 7-items
GHQ-28 = General Health Questionnaire 28-items
GHQ-9 = General Health Questionnaire 9-items
HADS = Hamilton Anxiety and Depression Scale
IES-R = Impact of Event Scale Revised
MASC = Multidimensional Anxiety Scale for Children
NDDI-E = Neurological Disorders Depression Inventory for Epilepsy
PCL-M = PTSD Checklist Military Version
PCL-5 = PTSD Checklist for the DSM-5
PDS = Post-traumatic Diagnostic Scale
PPQ = Perinatal PTSD Questionnaire
PHQ-9 = Patient Health Questionnaire 9-items
POMS = Profile of Mood States

## APPENDIX VI

## Results of Meta-regression for Anxiety

SYMPTOM	ANXIETY				
	Intervention Arm			Control Arm	
Variable	Regression Coefficient	p-value		Regression Coefficient	p-value
Sample Size	0.029	0.772		-0.088	0.270
Type of Journaling	-0.025	0.805		0.004	0.968
Study Duration	—			—	0.503
Journal Analysis	-0.155	0.057		-0.083	0.301
Sex	0.178	0.038*		0.181	0.002*
Age	0.017	0.871		-0.061	
Region-Europe	0.026	0.850	—	0.021	0.867
Region-North America	—			—	
Region-Australia	—			—	

## Results of Meta-regression for Depression

SYMPTOM	DEPRESSION				
	Intervention Arm			Control Arm	
Variable	Regression Coefficient	p-value		Regression Coefficient	p-value
Sample Size	0.024	0.616		-0.012	0.711
Type of Journaling	-0.050	0.226		-0.051	0.065
Study Duration	-0.104	0.005*		-0.056	0.035
Journal Analysis	-0.078	0.043		-0.012	0.678
Sex	0.041	0.299		0.025	0.357
Age	-0.003	0.939		-0.019	0.492

Region-Europe	-0.082	0.313	0.673	-0.024	0.684	0.722
Region-North America	-0.140	0.284		-0.108	0.295	
Region-Australia	-0.060	0.599		-0.042	0.610	

## Results of Meta-regression for PTSD

SYMPTOM	PTSD					
	Intervention Arm			Control Arm		
Variable	Regression Coefficient	p-value		Regression Coefficient	p-value	
Sample Size	-0.005	0.897		-0.001	0.980	
Type of Journaling	—			—		
Study Duration	—			—		
Journal Analysis	-0.017	0.647		0.006	0.803	
Sex	0.032	0.398		-0.012	0.642	
Age	0.073	0.012*		—		
Region-Europe	0.018	0.630	—	-0.013	0.555	—
Region-North America	—			—		
Region-Australia	—			—		

## APPENDIX VII

Results from the Cochrane risk of bias assessment tool for randomized control trials

<b>Study</b>	<b>Risk-of-bias judgement</b>	<b>Risk of bias judgement</b>				
	<b>Domain 1</b>	<b>Domain 2</b>	<b>Domain 3</b>	<b>Domain 4</b>	<b>Domain 5</b>	<b>overall</b>
<b>Alparone et al (2015)</b>	Some Concerns	High	Some Concerns	Low	Some Concerns	Some Concerns
<b>Bernard et al (2006)</b>	Low	Low	Low	Low	Some Concerns	Some Concerns
<b>Dennick et al (2015)</b>	Low	Low	Low	Low	Low	Low risk of bias
<b>Di Blasio et al (2015)</b>	Low	Low	Low	Low	High	Low risk of bias
<b>Ducasse et al (2018)</b>	Low	Low	Low	Low	Low	Low risk of bias
<b>Graf et al. (2008)</b>	Some Concerns	High	Low	Low	Low	High risk of bias
<b>Horsch et al (2016)</b>	Low	Low	Low	Low	Low	low risk of bias
<b>Jensen-Johansen et al (2012)</b>	Low	Low	Low	Low	Low	Low risk of bias
<b>Barton &amp; Jackson (2008)</b>	Low	Some Concerns	Low	Low	Low	Low risk of bias
<b>Koopman et al (2005)</b>	Some Concerns	Low	High	Low	Low	High risk of bias

<b>Barry &amp; Singer (2001)</b>	Low	Low	Low	Low	Low	Low risk of bias
<b>Lovell et al (2016)</b>	Low	Low	Low	Some Concerns	Low	Low risk of bias
<b>Suhr, Risch, &amp; Wilz (2017)</b>	Low	Low	Low	Low	Low	Low risk of bias
<b>Martino et al (2012)</b>	High	High	Low	Some Concerns	Low	High risk of bias
<b>Meshberg-Cohen et al (2014)</b>	Some Concerns	Low	Low	Low	Low	Some Concerns
<b>Possemato et al (2010)</b>	Low	Low	Low	Low	Low	Low risk of bias
<b>Rabiepoor et al (2019)</b>	Low	Low	Low	Low	Low	Low risk of bias
<b>Rawlings et al (2018)</b>	Low	Low	Some Concerns	Low	Low	Some Concerns
<b>Schache et al (2019)</b>	Low	High	Some Concerns	Low	Some Concerns	High risk of bias
<b>Wong &amp; Mak (2016)</b>	Low	Low	Low	Low	Low	Low risk of bias