1. INTRODUCTION

Diabetes Mellitus (DM) affects 422 million people globally (World Health Organisation, 2020a). Type 1 (T1DM) and Type 2 diabetes (T2DM) are its most common forms, together accounting for >95% of US cases (Xu et al., 2018). DM has profound personal and financial implications (Kristaningrum et al., 2021; Naylor et al., 2012). It is the world’s 8th leading cause of disability adjusted life years (Vos et al., 2020).Challenges reported by people with DM include depression and/or anxiety. US and UK estimates indicate people with DM are 2 to 3 times more likely to suffer from serious mental health conditions (Collins et al., 2009; Gendelman et al., 2009; Zghebi et al., 2020).

Depression and anxiety in people with DM are associated with reduced quality of life (Anaforoğlu et al., 2012; de Ornelas Maia et al., 2013), increased suicidal ideation (Elamoshy et al., 2018; Han et al., 2013), reduced treatment adherence, poorer glycaemic control and increased healthcare usage (de Ornelas Maia et al., 2013; Lin et al., 2004).

The high prevalence and impact of depression and anxiety highlight a need for effective interventions. A sizeable number of reviews have been completed in this area. Most conclude psychological and pharmacological interventions are efficacious (Baumeister et al., 2014; Chapman et al., 2015; Li et al., 2017; Uchendu & Blake, 2017; Van der Feltz-Cornelis et al., 2010). Accordingly, clinical guidelines recommend a range of psychological and pharmacological interventions for those with DM (Lipscombe et al., 2018; National Institute for Health and Care Excellence, 2021; RACPG, 2020; SIGN, 2017; Young-Hyman et al., 2016). What remains unknown is the actual clinical benefit of the interventions. We also lack a clear understanding of what accounts for the sizeable heterogeneity that exists between trials in the effects they found.

Little is known about clinical benefit because meta-analyses (MAs) have almost exclusively focused on effect sizes. Effect sizes are just one method of assessing change. They tell us about the statistical magnitude of the change elicited by an intervention at group-level. What they do not tell us is the clinical or practical relevance of that change – i.e., for what proportion of patients is the change sufficient to mean they recover (the optimal treatment outcome in mental health settings (Keller, 2003)).

To answer this question, an evaluation of clinical significance is required. Several methods exist. Comparisons of them show they yield similar results (Atkins et al., 2005). Jacobson's method – which requires access to individual participant data (IPD) - is the most widely used (Jacobson & Truax, 1992; Ogles et al., 2001; Wise, 2004).

According to Jacobson, two criteria need satisfying for clinically meaningful change to have occurred. Firstly, the difference between the person’s pre- and post-intervention distress score should be statistically reliable. A reliable change index (RCI) is used to determine whether the change exceeds that that which could be accounted for by known measure error. Secondly, the person’s post-intervention distress level needs to equate with that of someone from the 'functional' rather than 'dysfunctional' population. A cut-off score is calculated for the outcome measure being used to determine this.

Jacobson’s method means individuals can be categorised as: a) ‘recovered’ (i.e., shown statistically reliable change for the better and moved from a dysfunctional to a functional population); b) ‘improved’ (i.e., shown statistically reliable change for the better but not moved from a dysfunctional to a functional population); c) remaining ‘unchanged’ (i.e., made no statistically reliable change); or d) ‘deteriorated’ (i.e., statistically reliable change for the worse). Patients must be experiencing clinical levels of distress prior to intervention for the method to be used.

The approach has helped clarify the utility of interventions for distress in people with a range of other conditions, enabling stakeholders to make informed judgements about the efficacy of current interventions (Fisher et al., 2020; Noble et al., 2018; Temple et al., 2020). It is not though without potential challenges (e.g., (Noble et al., 2018)). For instance, it assumes data on trial participants’ distress is normally distributed (Jacobson & Revenstorf, 1988; Ronk et al., 2012; Tingey et al., 1996a). If this is not the case, the cut-off scores calculated to determine ‘recovery’ can become overly stringent (Temple et al., 2020).

To help further clarify the benefit of current interventions, it would be important to also consider potentially important moderators, which MAs to date have not. No MA has determined the importance of whether a trial used a manualised or non-manualized psychological intervention (Truijens et al., 2019), the relevance of dosage or length of drug treatment or the role of participants intervention adherence. Also neglected in most MAs is the potential importance of type (or format) of the psychological intervention and type of control or outcome measure used. MAs reliance on aggregate data also means they have not been able to determine what moderating role diabetes type has.

What has potentially further obscured the actual efficacy of available interventions is that previous MAs have relied on published aggregate. This means they have been unable to exclude data from participants within the original trials who were not experiencing clinical levels of distress immediately prior to intervention.

The present study sought to overcome the limitations of previous MAs by conducting an IPD meta-analysis (IPD-MA). We systematically identified randomised control trials (RCTs) evaluating psychological and pharmacological interventions for depression and/or anxiety in DM and assessed their methodological quality. IPD was sought from authors for participants in their trials with the aim of calculating the proportion of participants in each of the trials’ arms demonstrating ‘recovery’, ‘reliable improvement’, and ‘deterioration’ on the trials' primary outcome measures for depression and anxiety. Risk differences (RDs) compared the likelihood of participants recovering if they were in the trials’ intervention rather than control arm and moderators were explored. Traditional effect sizes for the trials were also calculated, using IPD where available.

**2. METHODS**

Reporting is according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Individual Participant Data guidelines (Stewart et al., 2015). The protocol was registered (Mather et al., 2019).

**2.1. Trial eligibility criteria**

Followed the Population, Intervention, Comparison, Outcomes and Study framework (Moher et al., 2009).

Whilst trials of interventions for anxiety were sought, only 4 were identified. Of these, 2 provided IPD, but Jacobson’s method could be applied to IPD from one (Supplementary File 1 for further details). As an IPD-MA for anxiety was not possible, the remainder of this report focuses on trials of interventions for depression.

***2.1.1 Population***

Trials were eligible if their participants had T1 or T2DM, comorbid depression and were aged ≥18 years. If inclusion criteria for age was absent from a trial report, mean participant age needed to be ≥18 years.

***2.1.2. Intervention***

Eligible psychological interventions were those using psychological techniques to alleviate symptoms of depression. No restrictions were placed on mode of delivery or who delivered it. Eligible pharmacological interventions used drugs appearing on the list of suggested antidepressants and anxiolytics to be used in Cochrane Reviews (CCDAN, 2019).

***2.1.3. Comparison***

Psychological intervention trials could have used ‘treatment as usual’, ‘wait-list’, or an ‘active’ control as their comparator condition. Only placebo controlled pharmacological trials were eligible.

***2.1.4. Outcomes***

Trials needed to use a validated self-report outcome measure for depression, and it should have been the primary outcome measure. To increase the number of eligible pharmacological trials, the latter criteria was relaxed for pharmacological trials so trials with depression as secondary outcomes were also eligible.

***2.1.5. Study***

Studies needed to be published in English in a peer-reviewed journal and use a RCT design.

**2.2. Search strategy, screening and data handling**

SCOPUS, MEDLINE, PsycINFO, PubMED and CLINHAL were searched from inception to December 2021. Reference sections of previous reviews were also hand-searched. Supplementary File 2 details search terms, how reviewers screened studies and extracted data. Also, described are the processes used to ensure only IPD from participants with clinical depression prior to intervention was used.

**2.3. Coding and scoring of potential moderators**

Trials were coded according to intervention type, format, manualization, control type, outcome measure and Risk of Bias (RoB) (Supplementary File 3). Participants were coded according to their diabetes type. It was not possible to code according to intervention adherence as most studies did not provide sufficient data for this.

**2.3. Statistical analysis**

***2.3.1 General analysis strategy***

Analysis focused on participants’ depression levels at the earliest assessment post-intervention within the trials.

***2.3.2 Comparison of eligible trials that did and did not provide IPD***

Trials were compared in terms of their sample size and participants’ age (Mann-Whitney-U) and intervention type and format (Chi-square). Standardised mean difference (SMD) effect sizes (Hedges' g) and 95% confidence intervals (95% CIs) were also calculated and the pooled SMD for trials that did and did not provide IPD were compared (Supplementary File 4).

***2.3.3 Clinical significance analysis***

As detailed in Supplementary File 5, an RCI was calculated for each outcome measure, as were cut-off points to indicate recovery had occurred. The proportion of participants within each trial arm classed as having ‘recovered’, ‘improved’, or ‘deteriorated’ was then calculated. Risk differences (RDs) compared the likelihood of participants having these different outcomes if they were in a trial’s intervention rather than control arm. Proportions with the different outcomes and RDs were pooled across trials using random-effects meta-analysis models (DerSimonian & Laird, 1986) (Mantel & Haenszel, 1959) and heterogeneity assessed (I2). Values >75% are considered substantial (Deeks et al., 2019)

The RD for improvement for Lamers et al. (2011) was a statistical outlier (its 95% CI was outside the CI of the overall mean RD). This could have been because it did not specify a minimum number of CBT sessions participants needed to receive (range 2 to 10). A sensitivity analyses was thus conducted excluding it.

***2.3.4 Effect size analysis***

SMDs of trials providing IPD were pooled using the inverse-variance random-effects model and heterogeneity assessed (Supplementary File 6). SMDs of 0.2 were considered, 0.5 medium and 0.8 large (Cohen, 1988).

Two trials had outlying effect sizes (their 95% CIs fell outside the overall effect size 95% CI). One was from Ebert et al. (2016) – the only trial testing a study-developed psychological intervention rather than an established one. A sensitivity analyses was conducted excluding it. There was no discernible reason for the outlying effect size for Ell et al.’s (2017) trial. No sensitivity analysis was therefore conducted excluding it.

***2.3.5 Moderators***

Sub-group analyses (Cochrane's Q-test) were completed to explore moderating effects of diabetes type, intervention type, format, manualization, control type, outcome measure and RoB on RDs and on effect sizes.

***2.3.4 Statistical software***

SPSS (V.27) was used for data curation and describing the sample; Comprehensive Meta-Analysis (V.3.3.070) for calculating RDs and effect-size IPD-MA; StatsDirect (V.3.1.22) for proportion IPD-MA.

**3. RESULTS**

***3.1 Study selection***

*3.1.1 Trials of psychological interventions and provision of IPD*

Database and hand searches identified 8,420 non-duplicate studies. Titles/abstracts indicated 8,296 were not eligible. Full text articles for the remaining 124 trials were screened. In total, 25 RCTs, published between 1998 and 2021, were identified as eligible (Supplementary File 7).

Twelve of the eligible RCTs provided IPD (total participants = 2070). The remaining 13 trials did not (total participants n = 917).

Of the n=2106 participants for whom IPD was provided, we included IPD from n=1385 (65.8%; (Intervention= 702; Control= 683) participants from the 12 trials in the depression IPD-MA. Exclusions occurred because: n=372 participants were found to have missing data for their pre-intervention assessment and/or first post-intervention assessment; n=282 did not have clinical depression pre-intervention; n=64 were participants from an ineligible trial arm of two trials having received an exercise intervention; and n=3 were not recorded as having either T1 or T2DM (Supplementary File 8).

*3.1.2 Trials of pharmacological interventions and provision of IPD*

Database and hand searches identified 1,587 non-duplicate studies. Titles/abstracts indicated 1,570 were not eligible. Full text articles for the remaining 17 trials were screened for eligibility. In total, 5 RCTs, published between 2000 and 2010, were eligible (Supplementary File 7).

Only 1 of the eligible RCTs (Echeverry et al., 2009) – which included n= 89 participants – provided IPD. The remaining 4 RCTs did not provide IPD (total participants n=157).

Due to the limited provision of IPD, no IPD-MA for pharmacological interventions could be completed and so no further results relating to pharmacological trials are presented in this report. Supplementary File 9 details the characteristics of the pharmacological trials, their findings and RoB scores.

***3.2 Sample and study characteristics***

*3.2.1 Trials of psychological interventions*

Of the 12 RCTs providing IPD, 4 were conducted in the Netherlands, 2 in Germany, 2 in the US, 1 in Australia, 1 in Croatia, 1 in Iran and 1 in the UK. The median age of their participants was 56 (range 18 to 91), 64.3% were female and 74.1% had T2DM.

Across the trials, a total of 13 different interventions and 12 control conditions were used. Eight of the interventions were Cognitive Behavioural Therapy (CBT) approaches, 2 Mindfulness-based Interventions (MBIs), 1 Acceptance Commitment Therapy (ACT), 1 Psychological Education (PsycEd) and 1 Therapist-Assisted (Supplementary File 3 for definitions).

Three of the trials used the Beck's Depression Inventory-I (BDI-I) as their outcome measure, 4 the Centre for Epidemiological Studies of Depression (CES-D), 3 the Patient Health Questionnaire-9 (PHQ-9), 1 the BDI-II, and 1 the Hospital Anxiety and Depression Scale (HADS). In most (66.6%) trials, the post-intervention assessment occurred 0 weeks after intervention completion, but across the trials it did range from 0 to 24 weeks.

Even after adjusting for the challenge of blinding participants to intervention allocation within a trial of a psychological intervention, the RoB scores showed 6 trials still scored as being at 'high risk’ of bias. Of the remaining trials, 2 were classed as of 'some concerns' and 4 'low risk’. Supplementary File 8 describes the trials further.

***3.3 Comparison of trials of psychological interventions that did and did not provide IPD***

*3.3.1 Characteristics of trials that did and did not provide IPD*

No significant differences existed between trials that did and did not provide IPD in participant age or interventions tested. The sample sizes of the trials providing IPD (median=140, range 40 to 348) were though significantly larger than those of trials not providing IPD (median=62.50 range 30 to 250) (Supplementary File 4). Also, the 13 trials not providing IPD were at higher RoB; 6 were at ‘high risk’ and 7 of ‘some concerns’ (Supplementary File 10).

*3.3.2 Effects sizes for interventions in trials that did and did not provide IPD*

The interventions tested within the trials providing IPD had significantly (p<0.001) smaller effects on depression (*g*=0.46, 95% CI 0.25 to 0.67) than the interventions within the trials that did not provide IPD (*g*=3.40, 95% CI 2.12 to 4.69) (Supplementary File 4). There was a substantial amount of heterogeneity across both trials that provided IPD (*I2=*80%) and those that did not (*I2=*98%). Heterogeneity between trials providing IPD and those that did not was also substantial (*I2=*95%).

***3.4 Clinical significance analysis for trials of psychological interventions***

*3.4.1 Recovery*

The pooled recovery rate from depression was 17% for treated participants and 9% for controls. The pooled RD was statistically significant (0.06, 95% CI 0.02 to 0.10), indicating 6% more treated participants recovered than controls. Between-study heterogeneity was moderate (*I2=* 47%). Supplementary File 11 shows the rates of recovery from depression in the intervention and control arms of the individual trials and the RDs.

*3.4.2 Improvement*

The pooled improvement rate for depression was 53% for treated participants and 31% for controls (Supplementary File 12). The pooled RD was statistically significant (0.22, 95% CI 0.10 to 0.34), indicating 22% more treated participants reliably improved than controls. However, there was a considerable amount of between-study heterogeneity (*I2=* 83%).

*3.4.3 Deterioration*

The pooled deterioration rate for depression was 3% for treated participants and 7% for controls (Supplementary File 13). The pooled RD was statistically significant (-0.03, 95% CI -0.05 to -0.01), indicating approximately 4% more control participants deteriorated than treated participants (*I2=* 0%).

*3.4.4 Influence of potential moderators on recovery, improvement, and deterioration rates*

3.4.4.1 Recovery

Diabetes type, intervention format, intervention manualization, control type, and RoB did not moderate recovery rates for depression (Supplementary File 11). However, intervention type (*Q*(1) = 4.06, *p*=.044) and outcome measure (*Q*(2) = 6.06, *p*=.048) did.

Specifically, individuals were more likely to recover than controls if they had been randomised to a MBI rather than CBT intervention. Of individuals randomised to an MBI, 28% recovered, with a RD of 0.20 (95% CI 0.06 to 0.33). In contrast, 17% of those randomised to a CBT intervention recovered, with a RD of 0.05 (95% CI, 0.01 to 0.10). Other intervention types (ACT, PsychoEd, TA) could not be included in the moderation analyses as there were <2 trials within each subgroup.

With respect to outcome measure, individuals randomised to an intervention arm were more likely to recover than controls if the CESD (RD=0.08, 95% CI 0.04 to 0.11) was used rather than the BDI (RD=0.01, 95% CI, -0.04 to 0.06) or the PHQ-9 (RD=-0.01, 95% CI -0.13 to 0.10). Other outcome measures (BDI-II, HADS-D) could not be tested as there were <2 trials in each subgroup.

3.4.4.2 Improvement

Diabetes type, intervention format, control type, outcome measure and manual use did not moderate improvement rates for depression (Supplementary File 12). However, intervention type (*Q*(1) = 3.90, *p*=.048) and RoB (*Q*(2) = 14.26, *p*=.001) did.

Intervention arm participants individuals were more likely to show improvement than controls if they were randomised to a MBI rather than CBT intervention. The RD for individuals randomised to an MBI was 0.34 (95% CI 0.12 to 0.56) with 44% improving. For those randomised to a CBT intervention the RD was only 0.20 (95% CI, 0.05 to 0.35). Other intervention types (ACT, PsychoEd, TA) could not be included in moderation analyses due to there being <2 trials in each subgroup.

With respect to RoB, those in the intervention groups were more likely to show improvement than controls if they were in a trial that had an ‘adjusted’ overall RoB classification of 'some concerns' (RD=0.43, 95% CI 0.33 to 0.54) rather than 'low' (RD=0.20, 95% CI -0.06 to 0.46) or 'high' (RD=0.15, 95% CI 0.05 to 0.25).

3.4.4.3 Deterioration

None of the variables examined moderated deterioration rates for depression (Supplementary File 13).

*3.4.5 Sensitivity analysis*

Having removed the RD for Lamers et al.’s trial, heterogeneity (I2) for improvement reduced from 83 to 71% and the pooled RD increased from 0.22 to 0.26 (95% CI 0.10 to 0.34). With the removal of the RD, whilst the moderating effect for RoB remained, intervention type was no longer a significant moderator (*Q*(1)= 3.02, *p*=.082).

***3.5 Effect size analysis for trials of psychological interventions***

A small-moderate, statistically significant effect in favour of the psychological interventions for depression was determined (*g*=0.46, 95% CI 0.25 to 0.67) (Supplementary File 6). However, there was considerable heterogeneity between trials in their effect (*I2*=80%).

*3.5.1 Influence of potential moderators on effect sizes*

Diabetes type moderated effect size, with effects being larger for patients with T1DM (*g*=0.77, 95% CI 0.48 to 1.05) compared to those with T2DM (*g*=0.35, 95% CI 0.31 to 0.59) (Supplementary File 6). Intervention type, format, manualization, control type and outcome measure did not moderate effect size.

*3.5.2 Sensitivity analysis*

When the effect size for Ebert et al.’s (2016) trial was removed, the size and significance of the effect in favour of psychological interventions remained, as did the moderating effect of diabetes type. Heterogeneity (I2) reduced marginally from 80 to 74%.

**4. DISCUSSION**

**4.1 Summary of results**

Psychological interventions are widely recommended for depression in people with T1 and T2DM. The extent to which they afford patients clinically relevant reductions in symptomatology was unknown. We therefore completed an IPD-MA and used Jacobson’s criteria to determine and compare the proportion of persons in intervention and control groups that actually recover, improve, and deteriorate.

Our findings are revealing. Whilst patients offered psychological intervention are statistically more likely than controls to recover, the proportion doing so is low. Across the trials, only ~17% of treated patients were found to recover. On average, just 6% more patients in the intervention groups recovered compared with controls. Despite potentially receiving up to 16 hours of intervention, the vast majority of treated patients did not recover and continued to experience clinical levels of depressive symptoms.

Thus, whilst currently available treatments are preferable to no treatment, our results emphasise substantial room exists for improving the efficacy of psychological interventions. Previous AD-MA have largely concluded that efficacious interventions exist, with all the potential implications this has for investment in refining existing interventions or developing new ones. Our findings are in line with those of recent IPD-MAs on interventions for psychological distress in other populations, including cancer and epilepsy (Fisher et al., 2020; Noble et al., 2018; Temple et al., 2020). They too found CBT elicits low rates of clinically meaningful change.

When psychological interventions in our IPD-MA were judged against the less stringent criteria of needing to show more improvement in depression than can be accounted for by measurement error alone, the picture was more favourable; 53% of intervention participants improved compared to 31% of controls. Intervention participants were also significantly less likely to deteriorate (3% vs 7%).

The findings from our effect size IPD-MA were also revealing. It generated much less striking effect sizes for psychological interventions than AD-MAs in the field. Specifically, psychological interventions were found to have a small to moderate effect on depression. Most previous AD-MA have reported psychological interventions have moderate to large favourable effects (Chapman et al., 2015; Li et al., 2017; Uchendu & Blake, 2017; Van der Feltz-Cornelis et al., 2010; Wang et al., 2017; Xie & Deng, 2017; Yang et al., 2020). The difference may reflect a difference in the trials we included; only 12 of 25 eligible studies provided IPD. However, it is possible that our results more accurately reflect true utility since previous MAs have completely relied on published data. It is known that AD-MAs and IPD-MAs can differ in their estimates of effect size for various reasons (Smith et al., 2016; Tierney et al., 2020). AD-MAs would not, for instance, be able to account for deterioration in the control group which could serve to exaggerate the apparent reduction in symptoms offered to participants from a psychological intervention being evaluated.

**4.2 Moderators of intervention effect**

As in previous MAs, there was sizeable heterogeneity between trials of psychological interventions. Our use of IPD permitted a different approach to exploring potential moderators.

When looking at moderators for recovery and improvement, factors such as diabetes type and control condition were not important. However, intervention typewas – participants receiving MBI interventions were more likely to recover or improve compared to controls than persons receiving CBT. We also found outcome measure moderated recovery, with participants being more likely to recover in trials using the CESD rather than the BDI or PHQ-9. In contrast, when effect sizes were examined, the only significant moderator to emerge was diabetes type, with effects being largest for participants with T1DM rather than T2DM.

The findings from these two sets of moderation analysis have different implications for clinical practice and research. One suggests UK guidelines recommendingCBT as the first-line treatment for distress in people with physical health conditions (National Institute for Health and Care Excellence, 2011, 2016, 2018) might require adjustment. The other suggests different treatment approaches might be required for people with T1 and T2DM. The conflicting results and different directions they have highlights the need for consensus on how to evaluate the efficacy of psychological interventions. For the reasons outlined in the introduction, we contend clinical significance is favourable.

**Strengths and limitations**

Whilst this study is the first of its kind and addresses many limitations of previous MAs, it is not without its own. Even after relaxing our inclusion criteria, only five eligible pharmacological intervention trials were identified, and only one provided IPD. This meant an investigation of the efficacy of pharmacological interventions was not possible. In addition, pharmacological trials were only eligible for inclusion if they used a drug that was on the Cochrane list. Thus, the efficacy of more novel pharmacological agents was not considered.

Further, due to the limited number of trials with a focus on anxiety, and limitations of those that did, our plan to determine the efficacy of psychological interventions for anxiety was not possible. More trials that focus on the treatment of anxiety are needed.

Also, we did not secure IPD for all 25 eligible trials of psychological interventions for depression; 50% provided IPD. Whilst this is favourable compared to previous IPD-MAs (Nevitt et al., 2017), it is unclear how inclusion of the IPD from these other trials might have changed our results. The trials not providing IPD only accounted for 31% of the 2,987 people with DM that had been randomised by the 25 trials altogether. We did though find those trials not providing IPD had larger effects in favour of interventions compared to those providing IPD for depression.

We were also unable to investigate the potential moderating effect of adherence, a key limitation in previous MAs in this area. Though IPD on adherence was requested from trial authors, the data was not provided for most. Moreover, due to inconsistencies in how adherence was reported in published articles it was not possible to use this. Future trials need to be supported to report adherence in a consistent manner.

Finally, it is important to understand the long-term effects of any psychological intervention. We were unable to investigate this due to inconsistent timings of follow-up assessments across eligible trials; they ranged from 1 to 12 months. This means the stability of the effects we found is unknown, Also, as intention-to-treat analyses could not be performed; we were only able to include participants who had completed baseline and post-intervention assessments, it is possible that the modest effects found immediately post-intervention were overestimated (McCoy, 2017), and outcomes could be somewhat worse than reported here.

**5. CONCLUSION**

Existing psychological interventions for depression offer limited benefit to people with T1 or T2DM, with less than 20% of treated patients recovering. Evidently, more efficacious interventions are required. It is possible that interventions that focus on psychological processes such as rumination may be more efficacious than interventions which focus on the content of cognition. High quality, controlled trials are needed to test this.

**Acknowledgements**

We would like to acknowledge all authors who contributed to the provision of IPD and thank them for the efforts involved.

We would like to acknowledge Sarah Dawson at Cochrane for her advice on the search strategy.

Finally, we would also like to acknowledge Dr James Temple for his comments on our findings and an earlier version of the manuscript.

References

Anaforoğlu, İ., Ramazanoğulları, İ., Algün, E., & Kutanis, R. (2012). Depression, anxiety and quality of life of family caregivers of patients with type 2 diabetes. *Medical Principles and Practice, 21*(4), 360-365.

Atkins, D. C., Bedics, J. D., McGlinchey, J. B., & Beauchaine, T. P. (2005). Assessing clinical significance: Does it matter which method we use? *Journal of consulting and clinical psychology, 73*(5), 982.

Baumeister, H., Hutter, N., & Bengel, J. (2014). Psychological and pharmacological interventions for depression in patients with diabetes mellitus: an abridged Cochrane review. *Diabetic Medicine, 31*(7), 773-786.

CCDAN. (2019). *CCDAN suggested classification of antidepressants for use in Cochrane reviews*. <https://cmd.cochrane.org/sites/cmd.cochrane.org/files/public/uploads/AD%20Classification.pdf></div>

Chapman, A., Liu, S., Merkouris, S., Enticott, J. C., Yang, H., Browning, C. J., & Thomas, S. A. (2015). Psychological interventions for the management of glycemic and psychological outcomes of type 2 diabetes mellitus in China: A systematic review and meta-analyses of randomized controlled trials. *Frontiers in public health, 3*, 252.

Cohen, J. (1988). Statistical power analysis for the social sciences.

Collins, M. M., Corcoran, P., & Perry, I. J. (2009). Anxiety and depression symptoms in patients with diabetes. *Diabetic Medicine, 26*(2), 153-161.

de Ornelas Maia, A. C. C., de Azevedo Braga, A., Paes, F., Machado, S., Carta, M. G., Nardi, A. E., & Silva, A. C. (2013). Comorbidity of depression and anxiety: association with poor quality of life in type 1 and 2 diabetic patients. *Clinical practice and epidemiology in mental health: CP & EMH, 9*, 136.

Deeks, J. J., Higgins, J. P., Altman, D. G., & Group, C. S. M. (2019). Analysing data and undertaking meta‐analyses. *Cochrane handbook for systematic reviews of interventions*, 241-284.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled clinical trials, 7*(3), 177-188.

Echeverry, D., Duran, P., Bonds, C., Lee, M., & Davidson, M. B. (2009). Effect of pharmacological treatment of depression on A1C and quality of life in low-income Hispanics and African Americans with diabetes: a randomized, double-blind, placebo-controlled trial. *Diabetes care, 32*(12), 2156-2160.

Elamoshy, R., Bird, Y., Thorpe, L., & Moraros, J. (2018). Risk of depression and suicidality among diabetic patients: a systematic review and meta-analysis. *Journal of clinical medicine, 7*(11), 445.

Fisher, P., Cherry, M., Stuart, T., Rigby, J., & Temple, J. (2020). People with obsessive-compulsive disorder often remain symptomatic following psychological treatment: A clinical significance analysis of manualised psychological interventions. *Journal of Affective Disorders, 275*, 94-108.

Gendelman, N., Snell-Bergeon, J. K., McFann, K., Kinney, G., Wadwa, R. P., Bishop, F., ..., & Maahs, D. M. (2009). Prevalence and correlates of depression in individuals with and without type 1 diabetes. *Diabetes care, 32*(4), 575-579.

Han, S. J., Kim, H. J., Choi, Y. J., Lee, K. W., & Kim, D. J. (2013). Increased risk of suicidal ideation in Korean adults with both diabetes and depression. *Diabetes research and clinical practice, 101*(3), 14-17.

Jacobson, N. S., & Revenstorf, D. (1988). Statistics for assessing the clinical significance of psychotherapy techniques: Issues, problems, and new developments. *Behavioral Assessment*.

Jacobson, N. S., & Truax, P. (1992). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research.

Kristaningrum, N. D., Ramadhani, D. A., & Yati Sri Hayati, S. (2021). Correlation between the burden of family caregivers and health status of people with diabetes mellitus. *Journal of Public Health Research, 10*(2).

Li, C., Xu, D., Hu, M., Tan, Y., Zhang, P., Li, G., & Chen, L. (2017). A systematic review and meta-analysis of randomized controlled trials of cognitive behavior therapy for patients with diabetes and depression. *Journal of psychosomatic research*(95), 44-54.

Lin, E. H., Katon, W., Von Korff, M., Rutter, C., Simon, G. E., Oliver, M., & Young, B. (2004). Relationship of depression and diabetes self-care, medication adherence, and preventive care. *Diabetes Care, 27*(9), 2154-2160.

Lipscombe, L., Booth, G., Butalia, S., Dasgupta, K., Eurich, D. T., Goldenberg, R., Khan, N., MacCallum, L., Shah, B. R., & Simpson, S. (2018). Pharmacologic glycemic management of type 2 diabetes in adults. *Canadian journal of diabetes, 42*, S88-S103.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the national cancer institute, 22*(4), 719-748.

Mather, S., Noble, A., Cherry, G., Fisher, P., & Nevitt, S. (2019). *A systematic review and individual participant data meta-analysis: The clinical efficacy of psychological and pharmacological treatments for anxiety and depression in adults with type 1 and type 2 diabetes* <https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019123180>

McCoy, C. E. (2017). Understanding the intention-to-treat principle in randomized controlled trials. *Western Journal of Emergency Medicine, 18*(6), 1075.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine, 6*(7), e1000097.

National Institute for Health and Care Excellence. (2011). *Generalised anxiety disorder and panic disorder in adults: Management. Clinical guideline 113*. <https://www.nice.org.uk/guidance/cg113/chapter/1-Guidance#principles-of-care-for-people-with-generalised-anxiety-disorder-gad>

National Institute for Health and Care Excellence. (2016). *Depression in adults with a chronic health problem. National clinical practice guideline 91*. <https://www.nice.org.uk/guidance/cg91/evidence/full-guideline-243876061>

National Institute for Health and Care Excellence. (2018). *Depression in adults: Treatment and management. NICE guideline: Short version*. <https://www.nice.org.uk/guidance/gid-cgwave0725/documents/short-version-of-draft-guideline>

National Institute for Health and Care Excellence. (2021). *Type 1 diabetes in adults: Diagnosis and management.* <https://www.nice.org.uk/guidance/ng17/chapter/Recommendations#managing-complications>

Naylor, C., Parsonage, M., McDaid, D., Knapp, M., Fossey, M., & Galea, A. (2012). Long-term conditions and mental health: The cost of co-morbidities. London: The King's Fund and Centre for Mental Health.

GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*, 1204-1222.

Nevitt, S. J., Marson, A. G., Davie, B., Reynolds, S., Williams, L., & Smith, C. T. (2017). Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: systematic review. *Bmj, 357*.

Noble, A. J., Reilly, J., Temple, J., & Fisher, P. L. (2018). Cognitive-behavioural therapy does not meaningfully reduce depression in most people with epilepsy: a systematic review of clinically reliable improvement. *Journal of Neurology, Neurosurgery & Psychiatry, 89*(11), 1129-1137.

Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical psychology review, 21*(3), 421-446.

[Record #109 is using a reference type undefined in this output style.]

Ronk, F. R., Hooke, G. R., & Page, A. C. (2012). How consistent are clinical significance classifications when calculation methods and outcome measures differ? *Clinical Psychology: science and practice, 19*(2), 167-179.

SIGN, S. I. G. N. (2017). *Management of diabetes: Quick reference guide.* <https://www.sign.ac.uk/assets/qrg116.pdf>

Smith, C. T., Marcucci, M., Nolan, S. J., Iorio, A., Sudell, M., Riley, R., Rovers, M. M., & Williamson, P. R. (2016). Individual participant data meta‐analyses compared with meta‐analyses based on aggregate data. *Cochrane Database of Systematic Reviews*(9).

Stewart, L. A., Clarke, M., Rovers, M., Riley, R. D., Simmonds, M., Stewart, G., & Tierney, J. F. (2015). Preferred reporting items for a systematic review and meta-analysis of individual participant data: the PRISMA-IPD statement. *Jama, 313*(16), 1657-1665.

Temple, J., Salmon, P., Smith, C. T., Huntley, C. D., Byrne, A., & Fisher, P. L. (2020). The questionable efficacy of manualized psychological treatments for distressed breast cancer patients: an individual patient data meta-analysis. *Clinical psychology review, 80*, 101883.

Tierney, J. F., Fisher, D. J., Burdett, S., Stewart, L. A., & Parmar, M. K. (2020). Comparison of aggregate and individual participant data approaches to meta-analysis of randomised trials: An observational study. *PLoS medicine, 17*(1), e1003019.

Tingey, R., Lambert, M., Burlingame, G., & Hansen, N. (1996a). Clinically significant change: Practical indicators for evaluating psychotherapy outcome. *Psychotherapy Research, 6*(2), 144-153.

Truijens, F., Zühlke‐van Hulzen, L., & Vanheule, S. (2019). To manualize, or not to manualize: Is that still the question? A systematic review of empirical evidence for manual superiority in psychological treatment. *Journal of Clinical Psychology, 75*(3), 329-343.

Uchendu, C., & Blake, H. (2017). Effectiveness of cognitive–behavioural therapy on glycaemic control and psychological outcomes in adults with diabetes mellitus: a systematic review and meta‐analysis of randomized controlled trials. *Diabetic Medicine, 34*(3), 328-339.

Van der Feltz-Cornelis, C. M., Nuyen, J., Stoop, C., Chan, J., Jacobson, A. M., Katon, W., ..., & Sartorius, N. (2010). Effect of interventions for major depressive disorder and significant depressive symptoms in patients with diabetes mellitus: a systematic review and meta-analysis . *General hospital psychiatry, 32*(4), 380-395.

Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., & Abdelalim, A. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet, 396*(10258), 1204-1222.

Wang, J., Wu, X., Lai, W., Long, E., Zhang, X., Li, W., Zhu, Y., Chen, C., Zhong, X., Liu, Z., Wang, D., & Lin, H. (2017, 2017-08-01). Prevalence of depression and depressive symptoms among outpatients: a systematic review and meta-analysis. *BMJ open, 7*(8), e017173. <https://doi.org/10.1136/bmjopen-2017-017173>

Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of personality assessment, 82*(1), 50-59.

[Record #15 is using a reference type undefined in this output style.]

Xie, J., & Deng, W. (2017). Psychosocial intervention for patients with type 2 diabetes mellitus and comorbid depression: a meta-analysis of randomized controlled trials. *Neuropsychiatric disease and treatment, 13*, 2681.

Xu, G., Liu, B., Sun, Y., Du, Y., Snetselaar, L. G., Hu, F. B., & Bao, W. (2018). Prevalence of diagnosed type 1 and type 2 diabetes among US adults in 2016 and 2017: Population based study.

Yang, X., Li, Z., & Sun, J. (2020). Effects of Cognitive Behavioral Therapy–Based Intervention on Improving Glycaemic, Psychological, and Physiological Outcomes in Adult Patients With Diabetes Mellitus: A Meta-Analysis of Randomized Controlled Trials. *Frontiers in psychiatry, 11*, 711.

Young-Hyman, D., De Groot, M., Hill-Briggs, F., Gonzalez, J. S., Hood, K., & Peyrot, M. (2016). Psychosocial care for people with diabetes: a position statement of the American Diabetes Association. *Diabetes care, 39*(12), 2126-2140.

Zghebi, S. S., Steinke, D. T., Rutter, M. K., & Ashcroft, D. M. (2020). Eleven-year multimorbidity burden among 637 255 people with and without type 2 diabetes: a population-based study using primary care and linked hospitalisation data. *BMJ open, 10*(7).