Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

	Please read the checklist for submitting comments at the end of this form. We cannot accept forms that are not filled in correctly. We would like to hear your views on the draft recommendations presented in the short version and any comments you may have on the evidence presented in the full version. We would also welcome views on the Equality Impact Assessment. We would like to hear your views on these questions: 1. Which areas will have the biggest impact on practice and be challenging to implement? Please say for whom and why. 2. Would implementation of any of the draft recommendations have significant cost implications? 3. What would help users overcome any challenges? (For example, existing practical resources or national initiatives, or examples of good practice.) See section 3.9 of Developing NICE guidance: how to get involved for suggestions of general points to think about when commenting.
Organisation name – Stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank):	**British Association for Counselling and Psychotherapy**

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**N/A**
Name of commentator person completing form:	**Dr Hadyn Williams, Chief Executive**
Type	[office use only]

Comment number	**Document** (**full** version, **short** version or the **appendices**	**Page number** Or **'general'** for comments on the whole document	**Line number** Or **'general'** for comments on the whole document	**Comments** Insert each comment in a new row. Do not paste other tables into this table, because your comments could get lost – type directly into this table.
Example 1	Full	16	45	We are concerned that this recommendation may imply that ……..
Example 2	Full	16	45	Question 1: This recommendation will be a challenging change in practice because ……
Example 3	Full	16	45	Question 3: Our trust has had experience of implementing this approach and would be willing to submit its experiences to the NICE shared learning database. Contact…….
1	Full	General	General	BACP have prepared this response to the 2017 NICE consultation on the revised *Guideline for Depression in Adults: Treatment and Management*, in our role as a professional body for UK counsellors and psychotherapists. As the largest British professional body for those providing psychological

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				therapies and as laid out in our mission statement (https://www.bacp.co.uk/about_bacp/) we aim to campaign for the highest standards of care for those experiencing depression. Moreover, our responsibility to both our members and the British public means that we campaign for a range of treatments to be available through the NHS for those with depression. This commitment reflects the considerable evidence of broad equivalence between therapies for depression (Gyani, Shafran, Layard & Clark, 2013; Pybis, Saxon, Hill, & Barkham, 2017; Stiles, Barkham, Twigg, Mellor- Clark, & Cooper, 2006; Stiles, Barkham, Mellor-Clark, & Connell, 2008) but also the evidence that it is important to give clients choice about treatment options because doing so improve treatment outcomes (Lindhiem, Bennett, Trentacosta, & McLear, 2014; Williams et al., 2016). It is important to note that this means that BACP has a commitment to support choice for all evidence-based therapies and as such welcomes the recommendations in the draft Guideline for the three main modalities practiced in the UK, namely Cognitive Behavioural Therapy (CBT), Psychodynamic Psychotherapy, and what is termed in the Guideline 'Counselling'. This document however focusses predominantly on counselling.
2	Full	General	General	This document was prepared by members of the BACP Research Department and draws on feedback on the draft Guideline from senior counselling and psychotherapy academic researchers in the UK and beyond. The document also draws on two reviews by academic teams independent of both NICE and BACP that were specifically commissioned by BACP to review the network meta-analysis and the economic cost modelling that informed the revised Guideline.
3	Full	General	General	BACP welcomes the extension to the original consultation period. However, given the length and complexity of the consultation documents and the level of detail required to digest and interpret the analysis, plus the timing of the consultation period, falling at a time of year when many researchers take holiday, we consider that the extension was wholly insufficient to allow for a properly robust independent level of scrutiny of the proposed Guideline recommendations and the processes and evidence used to arrive at such.
4	Full	208	General	An example of where there has been insufficient time to allow for proper scrutiny would be section 7.3.47 of the draft Guideline, which refers to the development of a "hierarchy of depression scales" "based on GC expert advice"; this hierarchy led to the inclusion in the network meta-analysis of data related to some scales but not others. No information is given in the documentation about either the

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				rationale for the prioritising of some instruments over others or the impact of data 'lost' from the analyses; it is possible that the impact of these decisions on the findings of the analyses was considerable.
5	Full	201 - 203	General	Another such example of insufficient time for appropriate scrutiny can be found in the process described in section 7.2.1 in which cut-off scores to distinguish between 'more' and 'less' severe depression were arrived at; given the lack of a substantial literature base to inform the decision, a "practical approach" (p202) was developed specifically for this Guideline review. This process was utilised to distinguish two study populations and was thus foundational in the process through which Guideline recommendations for 'more' and 'less' severe depression was arrived at. As a brand new procedure, one central to the entire enterprise, it would be entirely appropriate to interrogate not only the process utilised but also to systematically examine the implications of different cut-offs on the final analyses. However, again the lack of time available through the consultation period has not allowed for this point to be properly explored.
6	Full	General	General	BACP maintains in the strongest possible terms that detailed scrutiny of not only the evidence but the methods utilised is critical because historically the NICE Guideline for depression in adults has been significantly influential in shaping service delivery, in particular in England. As described by Clark (2011), the NICE recommendations for depression from 2004 onwards contributed to the development and roll-out of the Improving Access to Psychological Therapies (IAPT) programme, which in England and Wales now provides the bulk of treatment for depression in primary care (Gyani, Pumphrey, Parker, Shafran, & Rose, 2012). One example of the impact of the revised 2009 Guidelines appears to have been the cutting of counselling jobs in the NHS, with IAPT workforce census data suggesting a 35% decline in the number of qualified counsellors working as high-intensity therapists between 2012 and 2015, in a period where the total IAPT workforce grew by almost 18% (IAPT Programme, 2013; NHS England & Health Education England, 2016). Workforce shifts that apparently follow revised NICE guidelines (e.g. counselling not being recommended as a first line treatment for depression) underline the importance of scrutinising guideline recommendations since a core assumption is that using 'best' evidence and guideline methodologies will lead to NICE recommendations that improve patient care. In view of this we would argue that NICE has failed to facilitate a rigorous response to the consultation and in doing so has

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				not acted in the best interests of the public.
	Full	General	General	**BACP would like to make the following critiquing comments on the methodology used to arrive at the revised guideline (please see individual comments below):**
7	Full	35	18 - 22	***Conceptualisation of depression:*** The Guideline is based on an assumption that, for practical purposes, depression can be treated as a discrete entity, with a strong emphasis on evidence from controlled studies of interventions for depression. While these studies serve an invaluable function in generating scientific knowledge and understanding of this condition, we argue that they are of less relevance in relation to policy and practice because few patients present with clear-cut depression, for example, it is highly co-morbid with anxiety (Kaufman & Charney, 2000). In most cases, depression is one element of a complex set of problems that may include anxiety, cultural minority status, relationship difficulties, a life of adversity and trauma, loss and bereavement, work stress, physical incapacity and illness, and other factors (Smith, Court, McLean et al., 2014). It is our view that effective help and support requires a capacity to acknowledge, and if possible address, all of these dimensions. While the Guideline makes some attempt to consider these issues, for example in the section on social deprivation, we are concerned that this perspective does not sufficiently inform the main recommendations and, therefore, that patients who present with such comorbidities will not receive appropriate or effective treatment.
8	Full	General	General	***Privileging of RCT evidence:*** BACP cautiously welcomes the decision to retain counselling as a treatment. However, we are concerned about the recommendation to include counselling only as a second line intervention behind CBT and Behavioural Activation and we consider that this decision is based upon the privileging of RCTs above other relevant forms of evidence. We contend that the consideration of RCTs without the inclusion of very large routine practice-based datasets such as the IAPT dataset does not constitute 'the best available evidence', is unfit for purpose, and does not follow NICE's own procedural manual that "other study designs (including observational, experimental or qualitative) may also be used to assess effectiveness or aspects of effectiveness" (NICE, 2014/2017; p.15). We strongly recommend that NICE review their methodology to allow for the

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				inclusion of findings based on large routine practice-based data sets. We consider that there can be insufficient confidence in the results of RCTs for adult depression conducted to date that compare CBT with another therapy because they likely lack sufficient statistical power (Cuijpers, 2016). Similarly, meta-analyses of RCTs focused on treatment of depression are also vulnerable to low power (Cuijpers, 2016). It is our view that trials require much greater statistical power and less bias to determine differential effectiveness, and that from the existing RCT data it is unclear whether one therapy for adult depression is more effective than another to an extent which is clinically relevant (Barkham et al., 2017). In our view, this crucial point undermines the credibility that can be placed on the NICE-generated meta-analytic analysis that has been used to generate the recommendations contained in the draft consultation.
9	Full	General	General	***Failure to include large standardised routine datasets:*** The issues with power in RCTs and meta-analytic syntheses of such create in our view mean that there is little justification for relying solely on trials data and dismissing evidence from large standardised routine datasets such as the IAPT dataset since the size, methods of collection and analysis of routine datasets merit their inclusion (Barkham et al, 2017). The IAPT dataset also represents a considerable financial investment of taxpayers' money. It is our view that such large datasets collected at taxpayers' expense deserve greater respect and consideration and we contend that this data should be used in order to complement data from RCTs. It is important further to note that the evidence from the IAPT dataset is that counselling is as effective as CBT as an intervention for depression (Barkham et al, 2017). Existing evidence from IAPT annual reports (NHS Digital, 2014, 2015, 2016) demonstrates that patient recovery rates have been virtually equivalent between CBT and counselling (Barkham et al., 2017). Research on different portions of the IAPT dataset in relation to the treatment of depression have reported comparable outcomes between CBT and counselling (Gyani et al, 2013; Pybis et al, 2017). Given this, it is our view that IAPT data now needs to be considered alongside evidence from trials to form a more complete and accurate assessment of the comparative effectiveness of psychological therapies.
10	Full	General	General	***Failure to consider therapist effects:*** We also consider that a major gap in the guideline is the absence of attention to therapist effects. There is a growing body of evidence to indicate that there exit major differences between therapist effects (Barkham, Lutz, Lambert, & Saxon, 2017) or site effects (Pybis et al, 2017) where there appear to be noticeable differences in patients' outcomes (Saxon &

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				Barkham, 2012), with some studies finding that these effects are greater than the difference between alternative models of therapy it is our view that the current guideline risks being read by GPs and other front-line practitioners in such a way that they will expect that all CBT therapists, counsellors or other therapists are equivalent in effectiveness when this is not supported by the evidence. Inclusion of consideration of therapist or site effects would have the benefit of promoting greater attending to on-the-ground evidence of effectiveness of specific therapists and clinics.
11	Full	15; 49; 50	General	***Guideline Committee membership:*** We note that the committee membership is broadly described in section 3.3 (p50) of the draft Guideline and that the group members are named on p15 of the same document. We point out however that no information is given about the specific professional allegiances of the members of the guideline group, such as which therapies and interventions they have been trained in, or which they research, train others in, and currently use/ recommend to patients. This information is necessary for transparency and, in our view, is vital in order that the work of the group can be properly scrutinised and assessed for possible bias (Munder, Brutsch, Leonhart, Gerger & Barth, 2013). What is termed 'researcher allegiance' is a known biasing factor in psychotherapy research and in our view it is something that NICE should be systematically considering and seeking to protect against. This is critical since, as stated in the draft Guideline, the committee have a particular role in facilitating conclusions to be drawn in areas where there is a lack of data or findings are inconclusive (section 3.1, p49); in other words it is where the evidence is weakest that the role of the (potentially biased) committee is strongest.

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

12	Full	50	General	***Consideration of service user voice in revised Guideline:*** We note the inclusion of service users in the development of the draft Guideline (section 3.3.2) however we argue that this has been insufficient to ensure that service user voices have been properly included in the draft Guideline. This is because NICE elected not to update the guidance derived from qualitative data in this review of the Guideline; rigorously reviewing and synthesizing (Timulak, 2009) qualitative studies on service user experiences of depression and depression treatment is, we would argue, the only empirically supported way to ensure that a broad range of service user experiences are incorporated into the Guideline. The failure to systematically include service user voices in this way is disappointing given that NHS England's business plan for 2016/17 sets out a commitment: "to make a genuine shift to place patients at the centre, shaping services around their preferences and involving them at all stages" (NHS England, 2016, p.49). NICE has a similar commitment (NICE Patient and Public Involvement Policy, 2017). In our view, it is particularly egregious that NICE did not revisit the qualitative evidence around treatment of depression because while NICE processes do not (currently) allow such data to be included in the final summative analyses that shape key recommendations, a number of researchers (Hill, Chui, & Baumann, 2013; Midgley, Ansaldo, & Target, 2014) argue that qualitative outcome studies should be included. This is because they "offer a significant challenge to assumptions about outcome that derive from mainstream quantitative research on this topic, in relation to two questions: how the outcome is conceptualised, and the overall effectiveness of therapy" (McLeod, 2013, p.65). Reviewing existing literature, McLeod suggested patients themselves conceptualise outcome much more broadly than in terms of symptom or behavioural change (Binder, Holgersen, & Nielsen, 2010). Typically patients acknowledge ways in which therapy has been helpful but also where it has failed, suggesting that quantitative outcome research may overstate therapeutic effectiveness. Qualitative studies can also help answer questions about patient experience and expectations of NHS services, including whether treatments are credible and acceptable to them, which have an impact on outcomes.
	Full	General	General	**The recommendations in the NICE draft Guideline for Depression in Adults were developed out of a network meta-analysis and subsequent economic analysis. However review by BACP has identified a number of significant issues in the conducted NMAs that in our view suggest that the analysis results should be treated with considerable caution. These issues are described in the following sections below:** - **Selection of studies for inclusion**

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				- **Consideration of bias** - **Homogeneity of study population** - **Classification of interventions** - **Outcome variables selected** - **Statistical homogeneity** - **Inconsistency** - **Transitivity** - **Judgements related to Rankings**
13	Full	92; 216	General	***Selection of studies for inclusion:*** The network meta-analysis and consequent economic modelling are based on a selected group or studies; the inclusion/exclusion of studies thus significantly shapes the findings. BACP notes that while the process for selection of studies is detailed in the draft Guideline (section 4.5.2; p92) that the process is unclear and could have been improved. Specifically, we are concerned about the fact that it is difficult to understand which studies have been included in the various analyses conducted. For example, from Table 44 (Section 7.4.1.2; p216) it is not possible to determine which studies have been included to generate the N of 406 for counselling studies for the analysis pertaining to outcome related to SMD. We also note that it appears that different groups of studies have been included in each separate analysis for both the network meta-analysis and the economic modelling however again it is not clear which studies have been included in which analyses. The implication of the lack of clarity about the included studies is that a core process in the NICE analysis is not transparent and not thus amenable to review. More broadly we note what appears to be an arbitrary approach to selecting 'counselling' studies for inclusion. As an example, there is a notable lack of overlap with the studies included in Barth et al. (2013) who also conducted a network meta-analysis of treatments for depression and who also included counselling as a treatment. For example while the NICE analysis seems to have included eleven studies which included 'counselling' as an intervention, the Barth et al. (2013) study includes 37 studies. One study considered by NICE for inclusion was Cooper (2003); included in the Barth et al. (2013) analysis it was excluded from the NICE analysis on the grounds that the analysis focussed on post-partum depression although we are not sure why this was considered grounds for exclusion.

				As another example, the meta-analysis of psychological treatments for depression in primary care by Linde et al. (2015a) included a RCT on counselling by Corney and Simpson (2005) which does not appear to have been considered (e.g. it is not on the 'included' or 'excluded' list); the same meta-analysis included an RCT by Scott and Freeman (1992) comparing medication, CBT, counselling and routine care, which was included in the NICE analysis but, as far as we can see, not included in the analyses for counselling. The same point can be made about the RCT by Rosso, Martini and Maina (2013) which is included in the NICE analysis but although it includes a comparison between brief psychodynamic therapy with medication versus 'brief supportive therapy' with medication it does not seem to have been included in the 'counselling' analyses. The Guideline draft description of the method (section 3.5) also does not state whether data extraction and assessment of methodological quality were performed by independent raters, which is an established method to obtain reliable data. The inclusion of trials into any meta-analytic study is clearly critical in influencing the findings and we argue here that the decisions around inclusion/exclusion of studies in this network meta-analytic study can be criticised. We contend that this suggests the importance of not placing undue importance on the analysis results, especially considering the different conclusions drawn about 'counselling' by three other major and recent meta-analytic studies (Barth et al., 2013; Cuijpers et al., 2012; Linde et al., 2015b). For information, we also note that there are two trials currently underway which will contribute important data to the question of therapeutic effectiveness of Humanistic interventions: (1) a RCT on Emotion Focussed Therapy for depression being conducted in Portugal; (2) and the UK PRaCTICTED trial (Saxon et al., 2017).
14	Full	54	General	***Consideration of bias:*** BACP notes the various efforts to manage risk of bias. The Guideline authors did not use the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system for rating the quality of evidence, "because GRADE was not developed with network meta-analysis in mind" (Section 7.4 and 7.5). Although it is true, at least two GRADE-based evaluation systems are available for network-meta-analysis (Salanti et al., 2014; Puhan et al.,

				2014). It is acknowledged that these systems are recent yet although the Guideline authors address important GRADE-related issues while rating the quality of evidence, the assessment of quality of evidence conducted falls short of what is required by the two referenced systems, particularly with regard to the assessment of direct and indirect evidence (along with their methodological quality) for each effect estimate as well as regarding ranking treatments. In addition, there does not seem to have been broader effort to systematically consider bias related to researcher allegiance (RA). This is problematic given evidence from a meta-meta-analysis that "across n=30 meta-analyses the RA–outcome association was r=.262, corresponding to a moderate effect size" (Munder, Brutsch, Leonhart, Gerger & Barth, 2013). This is important, the study authors argue, since the estimate of the impact of researcher allegiance is greater than the typical difference in effectiveness between different types of therapy. As the authors of a network meta-analysis focussed on depression state of their own results: "Because data on a comparison level like allegiance cannot be considered in network meta-analysis, it is likely that researcher preferences influence the treatment effects found in this study to some extent" (Barth et al., 2013; p11). BACP argues that similarly Researcher Allegiance also likely significantly biased the individual RCTs included in the network meta-analysis and thus the overall findings. The Guideline authors also performed a sensitivity analysis in which the treatment effect estimates were adjusted for bias assumed to be present in small studies (as a proxy for publication bias), as described in Section 7.3 and Appendix N. They estimated bias for comparisons of active treatments with controls (while assuming that no bias is present in the comparisons of active treatments), and adjusted the effect estimates to account for this bias. As reported in Section 7.4 and 7.5, these analyses generally did not change the main conclusions, but in some cases had a substantial impact (changing the results) and sometimes were hard to interpret (bias in the opposite to the expected direction). Although these analyses may provide some rudimentary help for appraising the evidence, they are in general rather simplistic. Adjusting for risk bias in individual trials was not attempted, even though considerable variation in the methodological quality of the included trials was observed (Section 7.4 and 7.5). Overall it is clear that the NMA has in several areas not properly accounted for or managed risk of bias; this inevitably reduces confidence in the findings of the analysis.

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

15	Full	274	General	***Homogeneity of study population:*** Using symptom severity for the definition of study populations is likely to have accounted for effect modifiers that are associated with symptom severity. However, patients participating in trials testing different interventions may differ regarding factors that are only weakly related to severity. Even if the authors of the Guideline state that "a number of trials included... have successfully recruited participants who are willing to be randomized to either pharmacological or psychological intervention and to either self-help to face-to-face treatment" (Section 7.4 and 7.5), it is important to notice that the relative number of these trials is rather low. Furthermore, empirical findings suggest that participating in psychological vs. pharmacological trials is associated with clinically relevant patient characteristics and design factors (e.g., treatment duration, a certain type of control treatment) that are likely to influence treatment effect estimates (Linde, Rucker, Schneider & Kriston, 2016). In conclusion, although stratification by symptom severity probably reduced heterogeneity in the investigated populations, it is unclear whether these populations can be considered sufficiently homogeneous. This is problematic since a central assumption of network meta-analysis is that the populations investigated in a network are clinically homogeneous.
16	Full	274	General	**Classification of interventions** ***Clinical heterogeneity of interventions:*** The considered interventions were allocated to classes (e.g., selective serotonin reuptake inhibitors, cognitive and cognitive behavioural classes) and inferences were drawn both regarding single interventions and classes. While the decision to include also interventions that are considered clinically unsuitable (in order to enlarge the evidence basis) corresponds to up-to-date standards, the decision how to define interventions and interventions classes remains immanently subjective (Kriston, 2013). As Linde et al. (2015a) state: "Because psychological treatments are considered complex interventions, grouping them can be performed along several dimensions and remains controversial" [see also Craig et al, 2008]. For example, treatment as usual (defined as an intervention in the Guideline) is likely to encompass a wide range of interventions and it is improbable that it is clinically as homogeneous as Cognitive Behavioural Analysis System of Psychotherapy or short term psychodynamic group therapy (both of which are also defined as interventions in the Guideline). The grouping of interventions into classes is also not straightforward. For example, some might consider computerized cognitive behavioural therapy with support as part of the class of cognitive and cognitive behavioural therapies instead of the class of self-help with support as done in the Guideline.

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				The nomenclature and the exclusive classification system of the Guideline (which was necessary for being able to perform network meta-analyses) thus structures clinical reality in one of several possible ways, without an empirical basis, based on theoretical and clinical considerations. Even if the definition and classification of interventions was approved by the Guideline Committee, of which members are likely to have attempted to create practically relevant and broadly useful categories, they rely on assumptions that may not be shared by everyone. Specific examples of potential disagreements are provided in the section above on 'Selection of studies for inclusion;' another example of the potential impact of clinical judgement in creating classes is provided by Linde et al. (2015b) whose network meta-analytic study focussed on treatment of depression included a class they labelled 'other approaches' which was comprised predominantly of counselling studies but included one RCT of psychoeducation. In summary, the NMA utilises categorisation of the included studies into classes but the judgement about class membership is necessarily subjective; it is thus entirely possible that different groupings would have resulted in different findings from the NMA. ***Treatment as usual (TAU):*** Besides being a per se clinically heterogeneous category, the inclusion of TAU raises further questions. First, what is considered "usual" depends on the context of the studies that used it as a comparator. For example, a "usual" depression treatment is likely to be different in the UK, US, and Germany. Second, it is somewhat surprising that no combined intervention category with TAU has been defined, although traditionally several studies in depression compare TAU with TAU that is enhanced by the intervention of interest. It remains unclear, how this issue has been dealt with in the network meta-analyses. Third, if TAU is to be interpreted as usual care, then the results on more severe depression are rather discouraging, showing TAU to be statistically significantly less effective than pill placebo regarding symptom reduction and response (Section 7.5). It is not easily comprehensible that these results may mean that usual care (by definition the most frequent intervention) for treating severe depression is not supported by evidence. In conclusion, the practice of using TAU as a comparator in the network meta-analyses can be questioned.

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

17	Full	208	General	**Outcome variables** ***Standardized mean difference*:** The main clinical outcome was standardized mean difference (SMD) of depressive symptom severity change from baseline to the end of treatment as measured by continuous scales. For calculating SMD for change scores, sample size, mean change, and the standard deviation (or standard error) of change for each investigated group are necessary. As described in Section 7.3, this information (particularly the standard deviation of change) was not always completely available in primary study reports; therefore information on change from baseline was estimated form other figures (baseline end endpoint mean scores, standard deviations, and number of individuals in each group; or number of individuals responding in each group). Section 17.2 describes the detailed methods of this approximation, including the fact that it relies on information regarding the correlation of baseline and end-of-treatment scores as well as regarding the relationship between standard deviations at baseline and follow-up. While calculations of the Guideline authors in data from studies reporting necessary information support reasonable estimates for the latter, the correlation of baseline and-of-treatment scores ranged from 0 to 0.88 (Section 17.2). Based on these inconclusive empirical findings, the Guideline authors decided to assume a 6 correlation of 0.50, when it was not reported or directly calculable. Although sensitivity analyses assuming a correlation of 0.30 did not change the results, this assumption deserves further attention. In the analysis of SMD of symptom change for less severe depression, treatment effects in 86 of the 106 trials had to be estimated with the approximation method described above (Section 7.4), while this approximation was performed in 53 of the 68 trials for more severe depression (Section 7.5). Thus, the majority of the trial effect estimates was approximated. Even if the correlation between baseline and end-of-treatment scores was varied in sensitivity analyses, it was still assumed to be the same in all trials that required approximation, even though it was shown by the authors themselves that this correlation varies strongly across trials. As far more data were reported on end-of treatment scores (including standard deviations), using SMD of symptom severity at the end of treatment would have clearly been a better choice (relying much less on approximation). ***Response:*** Due to missing information in trial reports, the analysis of response data also relied strongly on estimating response from other information, essentially with the same methods and limitations as described for the SMD of symptom change.

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

18	Full	210	General	**Statistical homogeneity** The researchers investigated the homogeneity of the NMA analyses which is important as is a key assumption of network meta-analysis. ***Between-trial heterogeneity:*** It was assumed, that the statistical between trial-heterogeneity (the variation of the effect estimates) is the same for all comparisons of interventions. Although it simplifies statistical modelling, empirical findings suggest that this assumption is very unlikely to hold (Turner et al., 2016; Rhodes, Turner & Higgins, 2015). In addition, in some of the network meta-analyses moderate to high between-trial heterogeneity was present as compared to the average heterogeneity in a large number of meta-analyses (Salanti et al., 2014; Turner et al., 2016; Rhodes, Turner & Higgins, 2015), precluding firm conclusions regarding treatment effect estimates like in any meta-analysis. ***Within-class heterogeneity:*** Treatment effect estimates within intervention classes were assumed to be distributed around a mean class effect with a certain amount of within-class heterogeneity (unfortunately, within-class heterogeneity estimates are not reported in the Guideline). Due to sparse data, the prior distribution for this within-class heterogeneity parameter was informed by expert opinion (the network meta-analyses were performed in a Bayesian framework, in which 5 estimates are the result of updating prior distributions by data, with usually using uninformative priors that weigh data far more strongly than the prior). This prior, which strongly determined the estimated within-class heterogeneity due to the low amount of data, is described in Section 17.2 for binary outcomes (e.g., response, remission), defined for the logarithmic odds ratio. However, the sample code used for analysis of metric data with standardized mean differences reported in Section 17.6 uses this prior as well, although standardized mean differences are scaled on a somewhat smaller scale than logarithmic odds ratios. This means that the priority within-class heterogeneity for metric outcomes was higher than for binary outcomes. In addition, the within-class heterogeneity was per definition positive, leading to somewhat confusing findings. For example, in the analysis of standardized mean differences in comparison to pill placebo in patients with less severe depression, citalopram, escitalopram, fluoxetine, sertraline all have rather precise estimates (with -0.59 and 0.10 being the lowest of the lower and the highest of the upper bounds of the four 95% credible intervals, respectively, see Appendix W), but the credible interval for selective serotonin reuptake inhibitors as a class (consisting only of the four aforementioned interventions) is

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				-0.74 to 0.28 (Table 44 in Section 7.4). Although one would expect that estimates become more precise with more information, they actually seem to become more imprecise. In the same analysis, exercise as an intervention shows a clear effect (with credible interval -0.57 to -0.11, but exercise as a class (consisting only of exercise as intervention) has a credible interval of -1.57 to 0.89. It is also difficult to interpret the fact that according to the analysis of standardized mean differences for more severe depression, none of the treatment classes shows a statistically significant effect against placebo (by "statistically significant" meaning that the 95% credible intervals do not include the value for zero effect; Table 50 in Section 7.5). Even if these phenomena are in part likely to be the consequences of the class models drawing individual intervention estimates towards a class mean (Section 7.3) and borrowing within-class heterogeneity estimates from other classes in some cases (Section 17.2), they remain deeply unintuitive.
19	Full	Chapter 17	General	**Inconsistency** The researchers investigated the homogeneity of the NMA analyses which is important as another key assumption of network meta-analysis ***Global inconsistency:*** Network meta-analysis requires the assumption that treatment effect estimates from different sources (particularly form direct comparative trials and from indirect comparisons) are sufficiently homogeneous (it can be seen as the generalization of the homogeneity assumption from effect estimates within comparisons in pairwise meta-analysis to effect estimates across comparisons in network meta-analysis). This was done primarily by comparing a so called inconsistency model (not assuming consistency) with the main model (assuming consistency) regarding fit to the data (Section 7.3 and 17.2). This fit was assessed by comparing the deviance information criterion (DIC), a fit index, of the models. In these comparisons, a DIC difference of at least 5 points was considered as an indicator of difference between the models regarding their fit to the data, with no meaningful difference indicating that the consistency assumption is supported by the data. However, it has been suggested that a difference of two or three points can already be of practical relevance (Spiegelhalter et al., 2002), thus the threshold used by the Guideline authors may have been two conservative (missing important inconsistency). Furthermore, it is known that the presence of between-trial heterogeneity impedes the detection of inconsistency (Dias et al, 2013), and in some of the network meta-analyses reported in the Guideline considerable heterogeneity was identified (see above).

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				Local inconsistency: A careful evaluation of inconsistency should include checking the agreement between direct and indirect evidence for each comparison of interest (Salanti et al., 2014). By definition, it is possible only for comparisons for which direct comparative trials have been performed. Although direct and indirect evidence was not contrasted explicitly in the Guideline, Appendix W reports direct effect estimates from inconsistency models. Although direct evidence is available only for a fraction of comparisons, some of them suggest that direct and indirect effect estimates disagree to a substantial extent (for example, in the analysis of SMD of symptom change for sertraline vs. waitlist in less severe depression, the median estimated effect was -0.60 from network meta-analysis with consistency assumption and -1.24 from direct evidence). A more thorough inspection and explanation of local inconsistency would have been desirable for the assessment of the reliability of the reported findings. ***Difficulties in testing inconsistency:*** Inconsistency cannot be assessed for comparisons without direct evidence (or only regarding different sources of indirect evidence, which is currently not common practice). However, a careful investigation of most networks that are depicted in the Guideline reveals that the decisive body of evidence consists of two weakly connected sub-networks: one testing pharmacological interventions and using placebo as control treatment, while another testing psychological or physical interventions with waitlist or TAU as control. Even if these sub-networks can be consistent for themselves, due to sparse comparisons between them an essential part of inconsistency (for example, for comparisons of pharmacological and psychological treatments) cannot be assessed empirically.
20	Full	Chapter 17	General	**Transitivity** The third key assumption of network meta-analysis is transitivity (sometimes termed similarity). Linde, Rücker, Schneider & Kriston (2016) define transitivity as the requirement that the included trials comparing partly different sets of treatments (i.e., having different designs) are sufficiently similar with regard to clinical and methodological characteristics (e.g., population and outcomes). This means mainly that populations and interventions should be similar across different comparisons and that each participant could be, at least theoretically, randomized to any of the investigated interventions. The study by Linde, Rücker, Schneider & Kriston (2016) is relevant because the authors conducted a network meta-analysis of both pharmacological and psychological trials for depression in primary care

				and discussed the outcomes of this NMA in comparison with NMAs conducted separately for medications and psychological interventions. The authors concluded that while their assessment of the transitivity of the separate NMAs was broadly acceptable, the assumption of transitivity for the joint analysis was questionable. Notably the authors did not identify substantial heterogeneity and only moderate inconsistency in the joint NMA. The authors cautioned that: "Reviewers might be tempted to conclude that their network estimates are valid, if neither heterogeneity nor inconsistency is seen, both of which can be at least in part investigated and quantified by existing statistical tools, thus suggesting comforting ''objectivity.'' Still, it is crucial to notice that transitivity can be violated even in a homogeneous and consistent network, and its assessment inevitably needs qualitative clinical and epidemiologic appraisal." Thus, the authors, state, it is important "to be extremely careful about the interpretation of the findings from a network meta-analysis if the transitivity assumption is implausible." It is the conclusion of BACP that similar caution should be exercised in the case of this NMA. Transitivity is difficult to assess empirically, and therefore in most cases careful epidemiological judgment is necessary, for example using the criteria of Salanti (2012). Based on the expertise of the Guideline Committee, we could assume that this judgment was made as well as possible however this is difficult to assess given the failure to discuss transitivity in the draft guideline. Further, recalling the discussion above on treatment preference as a possible direct or indirect effect moderator, on the clinical heterogeneity of active treatments and TAU, and on the presence of two sub-networks of primarily psychological and primarily pharmacological interventions, transitivity of the analyzed networks can certainly be questioned. As Linde, Rücker, Schneider & Kriston (2016) comment, the evaluation of transitivity in network meta-analysis requires clinical judgment that may be subjective, context dependent, and accompanied by uncertainty, and the practical interpretation of findings from a network meta-analysis with uncertainty regarding its assumptions is correspondingly difficult.
21	Full	Chapter 17; pg51	General	**Judgements related to rankings of treatments** In general, results regarding the ranking of treatment according to their efficacy (as compared to placebo) were strongly emphasized throughout the Guideline. As an example, on p51 of Chapter 17, counselling is described as "the lowest ranked active class" of interventions in the SMD outcome

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				analysis. For this ranking of treatments, median and mean ranks were used, although more informative individual and summary approaches to ranking are available, such as the graphical display of rankograms and the calculation of the surface under the cumulative ranking curve (SUCRA) (Salanti, Ades & Ioannidis, 2011). However, treatment ranking altogether should always be interpreted with due caution (Salanti, Ades & Ioannidis, 2011; Ioannidis, 2009). For example, imprecision of treatment effect estimates is frequently associated with good (low) ranks. This is impressively demonstrated by the fact that in most analyses, interventions and intervention classes with the best (lowest) median ranks were tested in only a handful of patients, (usually less than 100, not rarely less than even 20) (Chapter 17).
	Full	General	General	**BACP is concerned that although the economic analysis undertaken has been conducted rigorously, that it is based on underlying assumptions which may mask the true comparability of the cost effectiveness of the interventions included (please see individual comments below)**
22	Full	239; 831; 832; 834	General	The economic analysis is largely based on the network meta-analysis which is in itself based on a number of assumptions which may be flawed. The problems in the NMA are detailed in the comments above. In addition the fact that flaws in the NMA have repercussions for interpretation of the economic analysis is acknowledged at various points in the Guideline draft: *Results [from the Guideline economic analysis] need to be interpreted with caution due to the limited evidence base characterising some of the interventions assessed in the models and methodological limitations characterising some of the NMAs that were used to populate the economic analyses. (p239)* *The quality and limitations of RCTs considered in the NMAs have unavoidably impacted on the quality of the economic model clinical input parameters. For example, economic results may be have been affected by reporting and publication bias. (p831)* *In addition, two of the NMAs that informed the economic analysis, remission in completers in less severe depression and discontinuation in more severe depression, were characterised by inconsistency between direct and indirect evidence, and therefore their results should be interpreted with caution. The limitations characterising the data included in the NMAs and the NMA outputs informing the economic analyses should be considered when interpreting the cost*

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				effectiveness results. (p832) For counselling in particular the significant lack of data on which the analysis is based is highly problematic. For counselling, the data on response in completers for less severe depression comprised only N=73; for remission in completers N=59; and for more severe depression, the data on response in completers was based on N=101. And no remission in completers data was available for counselling, which borrowed data from IPT (N=62) (p832). In fact the chapter which summarises the economic modelling of cost effectiveness (chapter 14) ends with the statement: "Results need to be interpreted with caution due to the limited evidence base characterising some of the interventions assessed in the models and methodological limitations characterising some of the NMAs that were used to populate the economic analyses" (p834).
23	Full	290	General	The economic analysis is based on the assumption that all psychological therapies are delivered by practitioners who are on the same pay scale as a band 7 clinical psychologist. This is not correct, many counsellors and psychotherapists delivering psychological therapies at step 3 within IAPT services and more broadly within the NHS are working at band 6, which makes them considerably more cost effective than this analysis would suggest. The guidance does acknowledge that the relative cost effectiveness of individual psychological interventions *en masse* will increase if any of these interventions can be delivered by a band 5 psychological wellbeing practitioner (PWP). There is also an option in scenario modelling to have a practitioner who has a unit cost halfway between the PWP and the psychologist unit costs. However, the guidance does not acknowledge that the relative cost effectiveness of these interventions will also change if one or more of the interventions can be delivered by Band 5 (or 6) practitioners while other still need to be delivered by Band 7 psychologists. Using Band 5 costings alone for counselling and keeping the higher costs for the other interventions would clearly increase the case for counselling. BACP would argue that the hourly costs of counselling are systematically lower than those for other psychological interventions and that as a result the relative cost effectiveness of counselling is underestimated.

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

24	Full	733	General	The economic analysis is based on the assumption that all psychological therapies are delivered in an equal number of sessions (16 sessions) but the draft guideline acknowledges that there is evidence that counselling may be delivered in fewer sessions (8 sessions). Changing the number of sessions of individual psychological interventions delivered would have an impact on the economic model. At the moment in the deterministic analysis all interventions are assumed to have 16 sessions of 1 hour in length. In the probabilistic sensitivity analysis versions for both models the number of sessions for individual psychological interventions can be varied randomly between a maximum of 16 and a minimum of 5, (with an 80% of being between 10 and 16 and a 20% change of being between 5 and 9). This assumes that the number of sessions provided by the interventions can genuinely vary randomly across interventions in this way. However if there is more of a systematic difference in the number of sessions offered for counselling specifically then the cost effectiveness rankings shown in the modelling are likely misleading.
25	Full	246	General	Combining the impact of lower pay and fewer sessions for counselling would also improve relative cost effectiveness. While the GC "also noted that according to the guideline economic analysis the cost effectiveness of counselling improved when this was effectively delivered by therapists paid at Band 6 or when this was delivered in 8 sessions, and agreed that these scenarios tested in sensitivity analysis may comprise variations of clinical practice in some settings" (p246), this was not systematically examined. An independent researcher commissioned to review the economic analysis for BACP modelled this and, holding all else constant, reducing the number of counselling sessions from 16 to 8 would mean counselling then being in the top 10 interventions for less severe depression and if also having a lower cost, e.g. using Band 5 costs, could even be in the top 5 interventions using the Net Monetary Benefit Approach for ranking.
26	Full	288	General	**Inconsistent use of economic findings** The draft Guideline states that the GC used the results of economic modelling (cost effectiveness) as the main criterion for making recommendations and the NMA results on the SMD of depressive symptom scores outcome (ranking of interventions and relative effects versus pill placebo) as a

				secondary criterion. However, for severe depression the recommendations do not include the use of counselling because of uncertainty over the effectiveness evidence, yet in this case the economic findings suggested that counselling is cost effective compared to usual management. It is difficult to understand how the decision to exclude counselling as a recommended treatment for depression was arrived at given the claims about how decisions were made.
27	Full	823	General	**Inconsistency in reported rankings following sensitivity analysis in less severe depression model** In Chapter 14 there is a discussion of the impacts on relative cost effectiveness of varying the bands of psychological workers. Page 823 states that: "W*hen all psychological interventions were assumed to be delivered by a band 5 PWP, the intervention cost of individual high-intensity psychological interventions was reduced and their relative cost effectiveness increased, resulting in changes in ranking. According to this scenario, the order of interventions from the most to the least cost-effective in deterministic analysis was as follows: mirtazapine, CBT group, physical exercise programme, CBT individual, IPT combined with citalopram, BA, citalopram, psychoeducational group programme, cCBT with support, cCBT without or with minimal support, physical exercise programme combined with sertraline, coping with Depression course (group), counselling, IPT, short term PDPT combined with citalopram, short term PDPT, clinical management, CBT individual combined with citalopram"* As a check on the analysis, an independent researcher commissioned to review the economic analysis for BACP used the version of the economic model without bias correction for less severe depression and applied these adjustments and found a slightly different order, e.g. with Exercise ranked 4th instead of 3rd and Counselling ranked 10th instead of 13th: The full order would be Mirtazapine, CBT group, IPT + citalopram, Exercise, BA, CBT individual, Citalopram, Psychoeducational group programme, cCBT with support, Counselling, Exercise + sertraline, IPT, Coping with Depression course (group), cCBT, Short term psychodynamic psychotherapy +citalopram, Short term psychodynamic psychotherapy, clinical management, CBT individual + citalopram. This inconsistency should be checked before model and guidance are published. Such errors do not foster confidence in the analysis.

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

28	Full	General	General	In summary, the BACP review of the economic modelling of the cost effectiveness of interventions for the treatment of depression in adults suggests that there are some inconsistencies in the analysis and that two key assumptions in the analysis are inappropriate for counselling. This is combined with the fact that the economic analysis builds out of the NMA which is itself – and as acknowledged in the draft Guideline – flawed in a number of ways. Overall, our conclusion is that the results are likely misleading and that the cost effectiveness of counselling as an intervention for depression in adults is not appropriately represented.
29	Full	248	14 - 26	**The evidence to support Recommendation 7.4.5, that psychotherapy and counselling interventions should be based on depression-specific treatment manuals, needs to be made explicit**. Studies in this area have produced mixed results, with some studies supporting the use of manuals, and other studies showing no advantage compared to treatment as usual (Carroll & Nuro, 2002). This is important because few manuals incorporate responsiveness to patient preferences (Ahn & Wampold, 2001). The use of manualised treatment therefore has the potential to undermine the principle of patient choice. There is also evidence that patient choice does not reflect existing brand-name established therapies. Instead, patient preferences tend to reflect a heterogeneous set of factors.
30	Full	252	24 - 27	**The evidence for the recommendation that any counselling intervention should be one developed specifically for depression (7.4.6) should be made explicit.** This requirement is only specified for counselling and short-term psychodynamic psychotherapy but not for CBT and IPT. What is the rationale for this?
31	Full	248	14 - 26	**Recording sessions** BACP welcomes the recommendation that 'healthcare professionals delivering interventions for people with depression should receive regular high-quality supervision'. However we are unclear on what is meant by 'external audit' as a way to monitor and evaluate competence. Whilst we can also see the benefits of recording sessions in regards to training and supervision, we

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				believe that there remains the potential for issues to arise in the near future following the introduction of the impending GDPR legislation in May 2018. In addition we urge caution as there is a significant lack of research around the impact of recording on the client experience.
32	General	General	General	**Recommendations for research** The draft guideline identifies a number of suggestions for research (e.g. p324). Based on the arguments made herein BACP would argue that what is necessary is: 1) RCTs which utilise CBT as a comparator; specifically RCTs on Humanistic Therapies focussed on both mild to moderate and severe depression. 2) Qualitative outcome studies with service users that focus on their experience of treatment of depression in primary care in the UK, in particular studies focussed on looking at experience of different therapeutic modalities; qualitative synthesis studies in the same area. 3) Research which seeks to systematically examine the differential impact of depression treatment for different groups in the UK. 4) Studies which examine therapist effects. A related recommendation is that NICE should review their approach to guideline development in line with some of the criticisms made here.
				References Ahn, H.-n., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology, 48*(3), 251-257. Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S., Znoj, H., ... & Cuijpers, P. (2013). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS medicine*, *10*(5), e1001454.

				Barkham, M., Lutz, W., Lambert, M. J., & Saxon, D. (2017). Therapist effects, effective therapists, and the law of variability. In L. G. Castonguay and Hill, C. E. (Eds.), *How and why are some therapists better than others?: Understanding therapist effects*. Washington, DC: American Psychological Association Barkham, M.; Moller, N. P. & Pybis, J (2017) How should we evaluate research on counselling and the treatment of depression? A case study on how NICE's draft 2018 guideline considered what counts as best evidence. *Counselling and Psychotherapy Research* Brown, E., Moller, N., & Ramsey-Wade, C. (2013). Recording therapy sessions: What do clients and therapists really think?. *Counselling and Psychotherapy Research*, *13*(4), 254-262. Carroll, K. M and Nuro, K. F (2002). One size cannot fit all: a stage model for psychotherapy manual development. Clinical Psychology: Science and Practice, 9(4): 393-406. Corney, R., & Simpson, S. (2005). Thirty-six month outcome data from a trial of counselling with chronically depressed patients in a general practice setting. *Psychology and Psychotherapy: Theory, Research and Practice*, *78*(1), 127-138. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. (2008) Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ; 337:a1655. Cuijpers, P. (2016) Are all psychotherapies equally effective in the treatment of adult depression? The lack of statistical power of comparative outcome studies. *Evidence Based Mental Health, 19,* 39-42. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. Med Decis Mak 2013; 33: 641-656. Gyani, A., Pumphrey, N., Parker, H., Shafran, R., & Rose, S. (2012). Investigating the use of NICE guidelines and IAPT services in the treatment of depression. *Mental Health in Family Medicine*, *9*, 149-160.

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				Gyani, A.; Shafran, R.; Layard, R. & Clark, D. M. (2013) Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy, 51,* 597-606. Ioannidis JPA. Ranking antidepressants. Lancet. 2009; 373: 1759-1760; author reply 1761-1762. Kriston L. (2013) Dealing with clinical heterogeneity in meta-analysis. Assumptions, methods, interpretation. International Journal of Methods in Psychiatry Research, 22(1):1–15. Linde, K., Rücker, G., Schneider, A., & Kriston, L. (2016). Questionable assumptions hampered interpretation of a network meta-analysis of primary care depression treatments. *Journal of clinical epidemiology*, *71*, 86-96. Linde, K., Sigterman, K., Kriston, L., Rücker, G., Jamil, S., Meissner, K., & Schneider, A. (2015a). Effectiveness of Psychological Treatments for Depressive Disorders in Primary Care: Systematic Review and Meta-Analysis. *Annals of Family Medicine*, *13*(1), 56–68. http://doi.org/10.1370/afm.1719 Linde, K., Rücker, G., Sigterman, K., Jamil, S., Meissner, K., Schneider, A., & Kriston, L. (2015b). Comparative effectiveness of psychological treatments for depressive disorders in primary care: network meta-analysis. *BMC family practice*, *16*(1), 103. Lindhiem, O., Bennett, C. B., Trentacosta, C. J., & McLear, C. (2014). Client preferences affect treatment satisfaction, completion, and clinical outcome: A meta-analysis. *Clinical Psychology Review, 34, 506 - 517* Munder, T., Brütsch, O., Leonhart, R., Gerger, H., & Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: an overview of reviews. *Clinical Psychology Review*, *33*(4), 501-511. NHS Digital (2014). Psychological Therapies: Annual report on the use of IAPT services. England, 2013-2014. http://content.digital.nhs.uk/catalogue/PUB14899 NHS Digital (2015). Psychological Therapies: Annual report on the use of IAPT services. England, 2014-2015. http://content.digital.nhs.uk/catalogue/PUB19098 NHS Digital (2016). Psychological Therapies: Annual report on the use of IAPT services. England, 2015-

				2016 http://content.digital.nhs.uk/pubs/psycther1516 Pybis, J., Saxon, D., Hill, A., & Barkham, M. (2017). The comparative effectiveness and efficiency of cognitive behaviour therapy and counselling in the treatment of depression: Evidence from the 2nd UK national audit of psychological therapies. *BMC Psychiatry, 17,* 215. Rhodes KM, Turner RM, Higgins JPT. Predictive distributions were developed for the extent of heterogeneity in metaanalyses of continuous outcome data. J Clin Epidemiol 2015; 68: 52-60. 6. Rosso, G., Martini, B., & Maina, G. (2013). Brief dynamic therapy and depression severity: A single-blind, randomized study. *Journal of affective disorders*, *147*(1), 101-106. Salanti, G. (2012). Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research synthesis methods*, *3*(2), 80-97. Salanti G, Giovane CD, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network metaanalysis. PLoS One 2014; 9: e99682. Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multipletreatment meta-analysis: an overview and tutorial. J Clin Epidemiol 2011; 64: 163-171. 11. Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology, 80,* 535–546. Saxon, D., Ashley, K., Bishop-Edwards, L….(2017). A pragmatic randomised controlled trial assessing the non-inferiority of counselling for depression versus cognitive-behaviour therapy for patients in primary care meeting a diagnosis of moderate or severe depression (PRaCTICED): Study protocol for a randomised controlled trial, Trials, 18:93 Scott, A. I., & Freeman, C. P. (1992). Edinburgh primary care depression study: treatment outcome, patient satisfaction, and cost after 16 weeks. *Bmj*, *304*(6831), 883-887.

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

				Smith, D.J., Court, H., McLean, G., et al. (2014). Depression and multimorbidity: a cross-sectional study of 1, 751,841 patients in primary care. Journal of Clinical Psychiatry. 75(11): 1202 – 8. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc Ser B Stat Methodol 2002; 64: 583-639. 8. Stiles, W.B., Barkham, M., Twigg, E., Mellor-Clark, J., & Cooper, M. (2006). Effectiveness of cognitivebehavioural, person-centred, and psychodynamic therapies as practiced in UK National Health Service settings. *Psychological Medicine, 36,* 555-566. Stiles, W.B., Barkham, M., Mellor-Clark, J., & Connell, J. (2008). Effectiveness of cognitivebehavioural, person-centred, and psychodynamic therapies in UK primary care routine practice: Replication in a larger sample. *Psychological Medicine, 38,* 677-688. Timulak, L. (2009). Meta-analysis of qualitative studies: A tool for reviewing qualitative research findings in psychotherapy. *Psychotherapy Research*, *19*(4-5), 591-600. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. (2012). Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol; 41: 818-827. 5. Williams, R., Farquharson, L., Palmer, L., Bassett, P., Clarke, J., Clark, D. M., & Crawford, M. J. (2016). Patient preference in psychological treatment and associations with self-reported outcome: national cross-sectional survey in England and Wales. *BMC Psychiatry, 16:*4.

Insert extra rows as needed

Checklist for submitting comments

- Use this comment form and submit it as a Word document (not a PDF).
- Complete the disclosure about links with, or funding from, the tobacco industry.
- Include page and line number (not section number) of the text each comment is about.
- Combine all comments from your organisation into 1 response. We cannot accept more than 1 response from each organisation.

Please return to: DepressionInAdultsUpdate@nice.org.uk

Consultation on draft guideline – deadline for comments 5PM on **12 September 2017 email:** DepressionInAdultsUpdate@nice.org.uk

- Do not paste other tables into this table – type directly into the table.
- Underline and highlight any confidential information or other material that you do not wish to be made public.
- Do not include medical information about yourself or another person from which you or the person could be identified.
- Spell out any abbreviations you use
- For copyright reasons, comment forms do not include attachments such as research articles, letters or leaflets (for copyright reasons). We return comments forms that have attachments without reading them. The stakeholder may resubmit the form without attachments, but it must be received by the deadline.

You can see any guidance that we have produced on topics related to this guideline by checking NICE Pathways.

Note: We reserve the right to summarise and edit comments received during consultations, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during our consultations are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory Committees.