swiss medical board

Point-by-Point Response

Stakeholder Comments for Health Technology Assessment

Antidepressants and cognitive behavioural therapy interventions for major depressive disorder

commissioned by the Swiss Medical Board

Assessment Team

Epidemiology, Biostatistics and Prevention Institute (EBPI), University of Zurich

Institute of Pharmaceutical Medicine (ECPM), University of Basel

09 August 2021

Stakeholder Comments

Full name	**Markus Ziegler**
Function	Head Market & IPR
Organization	Interpharma

Comment number	Chapter	Comment	Suggested change	Response
General Comments				
1	9 “Overall Conclusion” and report itself	To ensure an appropriate treatment environment, antidepressants, CBT interventions and the combination of antidepressants and CBT interventions are required. The therapy has to be chosen on a patient individual level to secure the required therapeutic success. Treatment with antidepressants can be initiated immediately after diagnoses without any delay due to access hurdles. Therefore antidepressants play a central role in the management to treat depression in all treatment periods (acute, continuation and maintenance).	Comment: We agree that there is no hierarchy within the treatment options antidepressants, CBT interventions and the combination of antidepressants and CBT interventions. All options are highly relevant to secure an appropriate treatment for the patient.	Thank you for your comments and your insightful feedback.
2	report itself	It is described that effects in three treatment periods will be assessed: acute, continuation and maintenance (e.g. 4.5 “Timeframe”; page 25). The presented comparisons within the HTA do not reflect those three periods consistently. Sometimes results are presented for all three periods,	Please ensure that results will be presented in a comparable and consistent way.	We agree that description of the effects of treatments across the different time periods should be done consistently. However, as mentioned in section 4.5 (page 25), the evaluation and presentation of the results by the three periods applied only to the clinical efficacy part. We have adapted the results of the main clinical efficacy section as well as the executive summary to present the findings by time

Comment number	Chapter	Comment	Suggested change	Response
		sometimes not. Sometimes results are presented specifically per intervention, sometimes just in general. This approach increases the risk of bias and inconsistency.		period within each comparison, including when there were no studies available.
3	4.6 “Study Designs” and report itself (page 26 ff.) 11.1.3 “Study characteristics” (page 187ff.)	It is mentioned, that evidence from RCTs is included to evaluate the clinical effectiveness and safety of ADM and CBT. The included RCTs are differently blinded (e.g. DB, RB, OL, SB). Included RCTs are published since 1. January 1995. Therefore a comparison is methodological extremely questionable, because outcomes are driven by blinding and structure of medical environment when the studies have been conducted. The risk is high that comparisons are methodological in-appropriate and results are not reliable and robust.	Please conduct only methodological appropriate comparisons to avoid unreliable results or explain in detail why the calculated comparisons are valid as the analysis do not follow the accepted or well established methodology.	This is a very important point. We agree that a major limitation in the evaluation of CBT interventions is the lack of blinding in trials assessing CBT and any comparison of these treatments will bear these limitations (since patients and clinicians cannot be blinded to CBT interventions). In addition to potentially affecting the transitivity of the network, the lack of blinding in CBT trials would likely produce a bias in favor of CBT. The potential issues arising from the lack of blinding are also addressed in the risk of bias assessment. For these reasons, we tried to include in the analyses only studies that were as similar in their design as possible, and we considered the primary analyses to be the pairwise meta-analyses providing the direct comparisons rather than the network meta-analyses. This may not have been explicitly clear from our report. We have added the following to section B of Methods in the Executive Summary and section 5.1.5 in the clinical efficacy review: *“Due to the heterogeneity of the studies in terms of study design (including the lack of blinding in trials evaluating CBT interventions) and participant characteristics potentially affecting the validity of results of the network meta-analyses, we considered the pairwise meta-analysis to be the primary analysis and we interpreted our findings accordingly”.*

Comment number	Chapter	Comment	Suggested change	Response
				We have additionally addressed this issue in the limitations part of section 5.3.2.1 : *"First, both pairwise and network meta-analyses assume that the included trials are drawn from the same population. While we only included adults with MDD, our dataset still included some populations with different age groups and specific medical comorbidities. Nevertheless, we tried to ensure homogeneity by excluding trials focusing on specific populations in the network meta-analyses and by performing sensitivity analyses on the pairwise meta-analyses. Network meta-analyses also assume that the interventions provided across the different studies are similar in terms of conduct and rationale. However, there were different classes of ADMs and a wide range of CBTs included in the studies, and the group of TAU may have merged different treatments (in terms of nature and intensity). This may violate the underlying transitivity assumption of our network and comprise the validity of our results. Along the same line, there were substantial differences in the study designs and conducts. While double-blinded placebo-controlled trials are usually employed for the evaluation of ADM, it is not possible to employ the same design in CBT (psychotherapy, in general) trials. It is essentially impossible to blind clinicians and patients to CBT. The inclusion of non-blinded trials in our network may affect the robustness of the estimates and the clinical benefit of CBT may be overestimated. Additionally, the choice of a true placebo in CBT trials is tricky and consequently other control groups are often used, such as waiting list and treatment as usual. It has been suggested that different control conditions in CBT trials lead to substantially different estimates and waiting list*

Comment number	Chapter	Comment	Suggested change	Response
				comparator groups could overestimate the effect of CBT. Furthermore, placebo effects of both ADM and CBT are well reported in the literature. While it is possible to estimate the placebo effect of ADM, it is hard to dismantle the true CBT effect from placebo as well as other non-specific factors such as patient expectancies, patient-doctor relationship and therapeutic alliance. Unfortunately, we could not explore the impact of placebo or nocebo effects of the different treatments on study estimates." In the same sections, we have discussed that differences between the studies may also be related to changes in practices and study conduct over time: *"Furthermore, we included studies that were conducted between 1995 and 2020 and there may be differences between past and current practices in conducting and evaluating the CBT interventions. Along the same line, clinicians rendering the interventions in the different studies may have had different training or experiences. Thus, the applicability of our findings may be limited. Future works should reconsider inclusion criteria including the date of publication and scope of intervention, or evaluate the impact of such differences on treatment effects."* A short summary outlining the limitations of all analyses have been added to the executive summary as well.

Comment number	Chapter	Comment	Suggested change	Response
4	4.2 “Intervention Antidepressants” and report itself (page 23 ff.)	21 antidepressants have been identified. Drug specific data are presented at maximum for 15 antidepressants only. It is unclear, why available evidence was not identified if the search terms included articles published since 1. January 1995.	Please update the search terms to ensure that all available evidence is identified and included in the HTA and to avoid a selection bias.	The list of 21 antidepressants was used in the first step of the search strategy to identify relevant systematic reviews. The 15 antidepressants were those that were included after updating the search conducted by Cipriani et al. using the same search terms and which reflect the most commonly used antidepressants, as specified in the protocol. We have added this to the section 5.3.2: *“We used the same search strategies as used by the authors of the two reviews to ensure consistency, accepting that some of the search terms were incomplete (e.g., LU AA 21004 was not contained as a vortioxetine search term). While we are confident that our search strategy captured relevant studies well, we cannot exclude that some may have remained unidentified.”*
5	4.2 Interventions	Nefazodone (NEF) is no longer on the market in Switzerland.	Please adjust	NEF has been removed from the analyses and updated results/tables are included in the revised document.
6	S.197 Class level network meta-analysis (upper triangle refer	SARI (it is not clear which molecules this group conatins). In literature, Nefazodone is sometimes classified as SARI. However, nefazodone is mentioned here independently under NEF. Further down (specific comments), however, trazodone is classified as a "serotonin modulator". We think that Trazodone is classified as SARI.	Please adjust Which products are classified as SARI? What does SARI stand for?	We have updated the analyses after removing NEF (comment above).
Specific Comments				

Comment number	Chapter	Comment	Suggested change	Response
7	11.1.2.2 (page 184)	Keywords used for identification of Vortioxetine RCT in Embase are incomplete.	As in other data bases of this HTA the authors should use the keywords “Vortioxetine” and “LU AA 21004”. Otherwise it is a selection bias in this HTA.	Thank you for pointing this out. We agree with you that the search term for vortioxetine is incomplete; however, we based our search on the same search strategy of the systematic review by Cipriani et al. (to ensure consistency) and unfortunately they did not include the keyword “LU AA 21004” (comment 4).
8	11.1.1.1 (page 174)	Vortioxetine is listed in the antidepressant group MAOI.	Vortioxetine (N06AX26) should be listed in the group “other antidepressants”	Edited accordingly.
9	4.4.1 (page 24)	The clinical effectiveness outcomes of interest do not correspond to the recommended outcomes parameters in international guidelines	The outcomes should correspond to the well accepted and valid endpoints that are recommended by guidelines and health authorities	We are unclear about the exact clinical and safety outcomes in the international guidelines that the reviewer is referring to. The outcomes in this HTA were agreed on together with the involved experts and stakeholders during the scoping process. The included outcomes cover all important outcomes commonly evaluated by clinical trials and are relevant for both clinicians and patients in the longer-term treatment of depression.
10	4.4.2 (page 24)	The safety outcomes of interest do not correspond to the recommended outcomes parameters in international guidelines	The outcomes should correspond to the well accepted and valid endpoints that are recommended by guidelines and health authorities.	Please see response to comment 3.
11	4.5 (page 25)	There is no rationale to use the time horizon of 12 months to estimate the cumulative risk	The estimation regarding benefit and harm should be done in a differentiated way using the accepted treatment phases as it is described in the HTA (Figure 1, page 20; Kupfer et al.,1999).	As mentioned in section 6.1.5, the rationale of using the time horizon of 12 months to estimate the cumulative risk of the benefit and harms outcomes was based on the average duration of a maintenance-phase therapy.
12	4 (page 22)	The description of all available antidepressants in one group	Create subgroups of antidepressants to avoid	We agree that it would certainly be interesting to assess any differences among the newer and older

Comment number	Chapter	Comment	Suggested change	Response
		embraces a very heterogenous group of antidepressants	heterogeneity. For example separate in “old” and “new” antidepressants resp.	generation of antidepressants. However, the analyses conducted were determined a priori in the scoping protocol and additional comprehensive subgroup analyses would be outside the scope of this report at this stage. Furthermore, it is to be noted that most antidepressants assessed in the studies were second or third generation antidepressants and only two studies included first generation antidepressants. All necessary estimates to make judgments about the individually included antidepressants in the trials are found in the report. We nevertheless conducted a sensitivity analysis by excluding older antidepressants (first generation antidepressants) from the network meta-analyses of the primary outcomes. Among all the primary outcomes, first generation ADMs were evaluated only in a few studies assessing acceptability, with no change in results after excluding them (Appendix 11.1.5.3).
13	4.1 (page 22)	Population of patients are diagnosed with DSM.	DSM is one of the accepted diagnostic instruments. ICD should be included, too.	ICD has been added to the examples of validated instruments.
14	5.2.3 (page 34)	A pairwise meta-analysis was not possible due to heterogeneity of studies.	The cut-off value to decide not to perform a meta-analysis due to heterogeneity should be included in the HTA	We did not conduct a meta-analysis if the studies differed from each other in terms of study design, conduct and intervention. We have added this to the text: *“A pairwise meta-analysis was not possible due to the heterogeneity of studies (i.e., due to different study designs) and scarcity of data.”*
15	5.2.4.1 (page 70)	The results of meta-analysis are assessed as “substantial heterogeneity”.	The ranges of assessment of heterogeneity should be included in HTA.	We have now included the thresholds for interpreting heterogeneity (section 5.1.5.1): *“Heterogeneity was statistically quantified using the I^2 statistic with its 95% CI and interpreted as following:*

Comment number	Chapter	Comment	Suggested change	Response
				0% – 40%: low, 30% – 60%: moderate, 50% – 90%: substantial and 75% – 100%: considerable, while taking into account the magnitude and consistency of estimates across studies."
16	5.2.4.1 (page 72)	Definition of the outcome parameter "remission" is missing	Criteria of "Remisison" should be defined according accepted cut-off values of psychometric scales	For both remission and response outcomes, we have now added a summary table with the definitions of each (Table 12 and Table 13).
17	Appendix 11.1.1.1 List of available antidepressants by class and drug in Switzerland (page 174)	Serotonin modulator	Trazodone is a SARI (Serotonin Antagonists and Reuptake Inhibitor)	Table has been updated.
18	Appendix 11.1.1.1 List of available antidepressants by class and drug in Switzerland (page 175)	Trittico (Vifor)	Trittico and Trittico retard (OM Pharma Suisse AG) Company name to be adapted	Name has been adapted.

Full name	**Prof. Dr. med. Erich Seifritz (Vorstand), Dr. med. Fulvia Rota (Präsidentin)**
Function	ES, Professor of Psychiatry and Chair of the Department of Psychiatry, Psychotherapy und

Organization	Psychosomatics, University of Zürich; FR, Board Certified Specialized Physician of Psychiatry and Psychotherapy FMH On behalf of the SGPP (Schweizerische Gesellschaft für Psychiatrie und Psychotherapie)

Comment number	Chapter	Comment	Suggested change	Response
General Comments				
19		Thank you for this comprehensive report and the invitation for review and comments by the SGPP. Since the report in its current form includes significant flaws both methological and regarding interpretation, we ask you to show us and other reviewers the revised report.		Thank you for your comments and valuable feedback. An updated report reflecting stakeholder comments and suggestions is provided along with this point-by-point response document.
20		The major critical issue of this paper is that the methodological approach of direct comparison between ADM und CBT efficacy appears problematic. ADM efficacy is evaluated in double-blind randomized controlled trials (dbRCTs), which involve 3 important preconditions: 1. random assignment of patients to a tx condition, 2. verum is compared with placebo, and 3. tx is employed in a double-blinded manner.	We recommend to - critically discuss the strengths and weaknesses of the methodological approach - revise the terminology effectiveness vs efficiency - implement the consequences of this differentiation in the conclusion statements - critically discuss the fundamental issues that CBT studies can not be	We agree with you that the review and analyses process suffer from several methodological limitations resulting from the types of evidence available for ADM and CBT. • As mentioned in response to comment 3 above, we have critically revised the clinical review by discussing these important aspects extensively within sections 5.3.2.1 and 5.3.2.2: *"First, both pairwise and network meta-analyses assume that the included trials are drawn from the same population. While we only included adults with MDD, our dataset still included populations with different age groups and specific medical*

Comment number	Chapter	Comment	Suggested change	Response
		In contrast, dbRCTs can not be employed to evaluate CBT's efficacy because of the following reasons: 1. placebo does not exist, 2. blinding is impossible, both for the patients and the therapist, and 3. the here most often employed waiting list group as a comparison condition of the CBT is in fact a nocebo. A consequence of this uneven comparison (ADM vs placebo; CBT vs nocebo) is a bias towards overestimated efficacy of CBT compared to ADM. Based on the above considerations, the general conclusions of the report are partially not supported by the current study. The major conclusions (p 17) criticized here include: 1. it is stated that no differences between ADM and CBT effectiveness were found. Since efficacy rather than effectiveness but has been studied, conclusion about effectiveness are not supported by the data. 2. since the comparison conditions (placebo vs nocebo) differ in ADM vs CBT trials, there is a	carried out as dbRCTs, and that waiting list comparison studies favour the investigated tx (nocebo effect) - include major metaanalyses and review and discuss these papers in the report - analyze and discuss unwanted effects of CBT the same way as those of ADM - please make only conclusive data-supported statements, or, if you make not data-driven conclusions tag them as personal opinion - it appears mandatory that the report carries out the same analyses for ADM and CBT. If data are not available for such analyses it should be stated explicitly in the report - the executive summary should clearly state the methodological weaknesses and limitation of the current report and the studies undelying this report; it	*comorbidities. Nevertheless, we tried to ensure homogeneity by excluding trials focusing on specific populations in the network meta-analyses and by performing sensitivity analyses in pairwise meta-analyses. Network meta-analyses also assume that the interventions provided across the different studies are similar in terms of conduct and rationale. However, there were different classes of ADMs and a wide range of CBTs included in the studies, and the group of TAU may have merged different treatments (in terms of nature and intensity). This may violate the underlying transitivity assumption of our network and comprise the validity of our results. Along the same line, there were substantial differences in study designs and conduct. While double-blinded placebo-controlled trials are usually employed for the evaluation of ADM, it is not possible to employ the same design in CBT (psychotherapy, in general) trials, It is essentially impossible to blind clinicians and patients to CBT. The inclusion of non-blinded trials in our network may affect the robustness of the estimates and the clinical benefit of CBT may be overestimated. Additionally, the choice of a true placebo in CBT trials is tricky and consequently other control groups are often used, such as waiting list and treatment as usual. It has been suggested that different control conditions in CBT trials lead to substantially different estimates and waiting list comparator groups could overestimate the effect of CBT. Furthermore, placebo effects of both ADM and CBT are well reported in the literature. While it is possible to estimate the placebo effect of ADM, it is hard to dismantle the true CBT effect from placebo as well as other non-specific factors such as patient expectancies, patient-doctor relationship and therapeutic alliance. Unfortunately, we could not*

Comment number	Chapter	Comment	Suggested change	Response
		methodological bias reducing the possibility of a direct comparison. 3. current leading literature including major metaanalyses have not been considered in the report. 4. the report intends to provide a translation from studies into real world (effectiveness) and to estimate the health economic effects. We recommend to improve the plausibility of the effectiveness conclusions which represent the basis for health economic statements. 5. the statement that the risk benefit ratio of CBT is better than that of ADM is not supported by the data, because of above outlined methodological reasons. 6. It is stated that the combination of ADM and CBT is not more efficient than each single treatment modality alone. This is a contradiction to major metaanalytic evaluations in the literature. We recommend to revise this statement according to the available data and the current literature. 7. the statement on p 17 “When compared to ADMs, CBT interventions seemed to provide a greater clinical benefit for MDD patients.” is not supported by the data.	is not sufficient to describe it somewhere in the almost 300 pages, but not in the main conclusions and the executive summary - since many stakeholders served as reviewers we would suggest that the report acknowledges them appropriately - we strongly suggest to revise the report accordingly and resubmit it to the reviewers/stakeholders for further comments	*explore the impact of placebo or nocebo effects of the different treatments on study estimates…...”* and, *“……...Second, when enough data was present to obtain pooled results, interpretation of the findings was often limited by small sample sizes and methodological limitations. Across our primary outcomes, we judged most of the studies to be at high risk of bias and none to be at a low risk. The potential bias in these studies, especially in those assessing CBT, was due to limitations (mainly driven by the nature of CBT) and flaws in the design and outcome measurements weakening the validity of results where high risk bias trials may tend to overestimate the benefits and underestimate the harms. Consequently, the quality of the evidence base drawn from these studies was too low to draw any firm conclusions. Third, we only included RCTs in our analysis. RCTs generally include highly selected populations, and include patients from secondary or tertiary care settings who may have higher levels of depression severity and be at a higher risk of relapse than the general population. Importantly, studies often had an enrichment design where patients were expected to have responded favourably to prior treatment. Such designs as well as procedures applied in clinical trials do not necessarily reflect real world practice. Thus, how these findings apply to the broader clinical population, including patients with milder MDD forms and who were underrepresented in such trials, is unclear. A fourth limitation is the lack of or inadequate assessment of the adverse effects of treatments, particularly relating to CBT. This*

Comment number	Chapter	Comment	Suggested change	Response
		8. A chapter of adverse effects of CBT is missing. All efficienc tx have side effects, this should be discussed the same way as for ADM. The benefit-harm balance can only be calculated if the efficacy of ADM and CBT would be assessed based on the same methodological approach (double-blind, which is obviously not possible in CBT trials), as reported e.g. on p 92 and further on. Waiting lists comparison groups are nocebo. A true placebo control in CBT trials would include personal discussions and interactions between a therapist and patient with the same setting and duration as the CBT. Obviously, this has not been done in the studies used for the current report. Studies specifically comparing CBT vs lay tx or dismantling studies did not show advantages of CBT, this effect is called Dodo-effect and is due in part to expectation and other unspecific effects. The literature is rich on placebo effect studies in psychotherapy, however, no single paper has been cited in the report. Therefore the figures given in the report on "net-clinical" benefit are not supported by the data. In section 6.3.1. the report discusses limitations, however, the statements can not correct false, i.e.		*inconsistent reporting of AEs in CBT trials may lead to an underestimation of their harms….."* • We have further added to the executive summary a short paragraph on these limitations (section A in the results part). • We have replaced the term "effectiveness" with "efficacy" where appropriate, and briefly discussed the applicability of the findings to real-world practice (Executive summary conclusion, section 5.3.2.2). • Regarding describing the adverse effects of CBT, we have already described the available evidence on AE not only for ADMs but also for CBT, their combination and placebo in section 5.2.4.4. • We agree that a benefit harm assessment of ADM and CBT is limited by the unequal reporting of adverse effects in trials reporting CBTs. For that reason, we considered acceptability (i.e., study drop-outs for any reasons) as a harm outcome since it is more commonly reported in ADM and CBT trials. We found similar results when either acceptability or specific adverse effects were considered as a harm outcome. The following has been added to section 6.3.1: *"Additionally, as reported in the clinical review, a serious limitation of the evidence*

Comment number	Chapter	Comment	Suggested change	Response
		not data-based, statements in the previous sections. In chapter 6.4. the report makes comparative statements on harm risk of CBT and ADM. Since most CBT trials do not report side-effect, and all ADM dbRCTs are obliged to report them systematically, there is a strong bias in this comparison, i.e., the conclusion is not supported by the available data, and the methodological approach, respectively.		*base on treatment relates to inconsistent AE reporting, particularly in trials evaluating CBT, which may lead to an underestimation of its true harms. As such, we did not rely solely on specific AEs as our harm outcome but also alternatively used all-cause study drop-outs which is more consistently reported in studies, with similar results when using either harm outcome."* • The conclusions in the executive summary and clinical efficacy section, and the overall conclusion have been revised. • All involved stakeholders and experts involved in all the stages are acknowledged in the updated version.
21		On p 17 (Conclusions) and elsewhere the report makes statements about the comparative cost-effectiveness of the ADMs available in Switzerland. We do not agree that the conclusion is supported by the data.		In order to temper our findings with some caution, we have now deleted the sentence in the conclusion section regarding the comparative cost-effectiveness of the ADMs in Switzerland.
22		If cost-effectiveness of different ADMs are compared why does the report not compare the cost-effectiveness of different CBTs listed in Tab. 11.1.1.2.		We did already describe separately the cost-effectiveness findings for internet-based CBT and for face-to-face CBT (page 166).
23		In Chapter 9 (Overall conclusions), the report makes statements that are not supported by the underlying		The conclusion has been revised to include a statement on the applicability of the findings from

Comment number	Chapter	Comment	Suggested change	Response
		data. If statements are made about effectiveness, the report should consider that in Switzerland it is very unusual that psychiatrists just prescribe ADM without accompanying the patient with some sort of psychotherapy, this is fundamentally different from the situation in dbRCTs. This should be discussed, especially in a report that claims to analyze the situation in Switzerland.		trials to real-world clinical practice (executive summary and section 9): *"In conclusion, we found that while both ADM and CBT interventions appear to be efficient options in the management of MDD, none appeared to definitively perform better than the other. While our benefit-harm assessment seemed to show that CBT interventions may provide a greater clinical benefit for MDD patients compared to ADM (due to lower reported adverse effects in CBT), the systematic review of the cost-effectiveness analyses, adapted to the Swiss context, revealed that CBT may also be more expensive depending on how it is conducted (number of sessions, online vs. face-to-face). These findings need to be carefully interpreted given the methodological shortcomings of the review and the evidence base, as well as the potential limited applicability of findings of controlled trials to real-world clinical settings."*
24		On P 164 the report states "The available evidence on relative treatment effects, outcomes and preferences showed favourable benefit-harm balance for CBT over ADM and for combination of CBT and ADM over a monotherapy.". This statement is not supported by the report itself and the underlying data. Of course there are more unwanted effects recorded in studies where these unwanted effects are systematically evaluated and recorded (as in all ADM RCTs) compared to studies where these effects are not evaluated and not		Please see response to comment 20.

Comment number	Chapter	Comment	Suggested change	Response
		recorded (most CBT studies). It should be noted here that clear evidence in the literature shows that dropout rate is smaller in CBT compared to ADM studies, and that low dropout rate has an effect on effect size. The smaller dropout rate is clearly associated with study and tx setting differences in CBT vs ADM trials.		
25		In the general conclusion statement and elsewhere in the report it is stated that ADM studies were sponsored by pharmaceutical companies, whereas all of the cost-effectiveness analyses of CBT included in this review were public sector funded. In this statement of the report it is suggested that ADM studies are biased by sponsor, and CBT studies are objective. However, clear evidence shows that CBT study results correlate linearly and positively with the allegiance of the principal investigator to the CBT under study. Thus, here the report neglects significant literature.		Thank you for raising this important point. We did not mean to imply that CBT trials are objective while as ADM studies are biased by their sponsors. We agree that psychotherapy trials in general are vulnerable to researcher allegiance. However, as mentioned in section 5.3.2.2, we were unable to explore its effect on our results. We have removed the statement mentioned statement from the overall conclusion and we have added the following to the limitations in section 7.3: *"On the other hand, it is well known that treatment effects in CBT trials are often associated with researcher allegiance, implying a potential for bias (175–177). "*
26		In the very final statement of the report (p 165) statements are made that are not supported by the data and the literature, and also contradicts the previous sections.		Conclusion has been revised.
27		Finally, it remains unclear which claims in the report are Switzerland-specific, and where there might be differences to other countries.		In the clinical efficacy review, we included all international studies while in the cost-effectiveness analysis we included those which were relevant for the Swiss context, as commonly done in HTAs.

Comment number	Chapter	Comment	Suggested change	Response
				Statements on cost-effectiveness are necessarily jurisdiction-specific as they depend on local cost-structures and healthcare practice patterns.

Full name	**Marianne Eggenberger, Markus Gnägi**
Function	Project Manager Medicines, Head of Department Official Tariffs and HTA
Organization	santésuisse

Comment number	Chapter	Comment	Suggested change	Response
General Comments				
28	HTA overall	The HTA is well and clearly structured. The question "Medium- to long-term use of antidepressants in MDD" is comprehensively addressed and the results as well as the discussion and conclusion is presented in a detailed and comprehensible manner.		Thank you for your valuable feedback.
29	1	The summary is already very detailed. The various and most important findings from the HTA are well summarized, particularly with regard to the findings on clinical effectiveness (efficacy) and safety of the medium- to long-term use of ADM and / or CBT. The findings on costs are also adequately included.	-	No response required.

Comment number	Chapter	Comment	Suggested change	Response
30	4	The decision context is focused on specific aspects in various points. Since antidepressants are used for different indications, the restrictions are understandable. Thus, this HTA only addresses a "small" and specific part of the problem of the long-term use of antidepressants. Accordingly, the limits of the HTA must be taken into account, although this is well illustrated in chapter 5.3.		No response required.
31	5	The chapter on clinical effectiveness and safety is detailed and comprehensible.		No response required.
32	6	Conducting a benefit-harm analysis is very interesting. Even if the findings obtained from it should be viewed with caution, the analysis carried out certainly points in an interesting direction. For the planning of future clinical trials, the findings can also be helpful and should be taken into account		No response required.
33	7	The results of the health economic overview can be reproduced.		No response required.
34	8	The results of the budget impact analysis provides a good overview of the relevant cost drivers.		No response required.
Specific Comments				
35	4	In the course of the HTA, preplanned subgroups are mentioned at various positions. However, no details can be found on those.	It is proposed to add Information on these preplanned subgroups in Chapter 4, even if they could not be considered later in the HTA due to the lack of data.	Thank you for your comment. The subgroups which were planned are reported in section 5.1.7: *"We were unable to conduct the pre-planned subgroup analyses (i.e., for age groups, MDD severity, country settings (industrialized versus non-*

Comment number	Chapter	Comment	Suggested change	Response
				industrialized), dosing schedule of antidepressants, delivery format of CBT, and treatment setting) due to sparsity of stratified data reported in relevant RCTs."
36	5.1	The two-stage literature search is presented in a comprehensible and understandable way. An important conclusion of the HTA is that further clinical studies are necessary to evaluate the problem addressed here. We would therefore have expected that a search in registers of ongoing studies would also have been conducted.	In future HTAs, it is proposed to include clinical trial registries in the literature search.	Thank you for your comment. We agree that the inclusion of clinical trial registries is important. The search through the Cochrane database automatically captures trials registered in WHO's International Clinical Trials Registry Platform and clinicaltrials.gov. However, of course, we cannot rule out that we may have missed other trials which were registered elsewhere. We have added the following to section 5.3.2.1: *"In addition, we were unable to contact authors of studies if there was missing or unclear information. We also did not search for grey or unpublished literature. However, we are fairly certain that our search through the Cochrane database captured RCTs in trial registries well enough."*
37	5.1	RCTs were included. Current guidelines are hardly mentioned. Especially in a topic where, in addition to the use of antidepressants, non-drug therapies are also used and accordingly serve as comparators, it would be desirable to include currently relevant guidelines and place them in the context found.	In future HTAs, where especially non-drug therapies are included in the evaluation, the topic of guidelines should be adequately addressed.	That is an important point, and we agree that the interpretation of the findings needs to be addressed within the context of current guidelines. However, it is to be noted that in current guidelines in Switzerland (as well as in Europe in general) the choice of initial and continued/maintained treatment depends on the severity of MDD. For example, in patients with mild depression, the Swiss guidelines recommend psychotherapy or pharmacotherapy after a period of watchful waiting (1) Edith Holsboer-Trachsler med, Josef Hättenschwiler M, med Johannes Beck P, phil Serge Brand P, med phil Ulrich Michael Hemmeter P, Martin Ekkehard Keck M, et al. Erhaltungstherapie und Rezidiv- prophylaxe unipolarer depressiver Störungen. 2016.,2) Holsboer-Trachsler E, Hättenschwiler JA, Beck J, Brand S, Hemmeter UM, Keck ME, et al. Die Akutbehandlung depressiver Episoden. Swiss Med Forum –

Comment number	Chapter	Comment	Suggested change	Response
				Schweizerisches Medizin-Forum. 2016 Aug 30;16(35), 3) https://www.medix.ch/wissen/guidelines/psychische-krankheiten/depression/). However, we were unable to explore whether there were any differential effects across the different severities of MDD due to the scarcity of the data; and thus, interpreting our findings in the context of current recommendations would be difficult. The following has been added to section 5.3.2.2: *"Importantly, the data offers no insights on how treatments might differ based on patients' severity of depression. This limits the interpretation of our findings in the context of current clinical guidelines where the recommended initial and maintained treatment strategies are based on severity of depression symptoms."*

Full name	**Egemen Savaskan / Dan Georgescu**
Function	Project Manager Medicines, Head of Department Official Tariffs and HTA
Organization	Klinik für Alterspsychiatrie, Psychiatrische Universitätsklinik Zürich / Klinik für Konsiliar-, Alters- und Neuropsychiatrie, Psychiatrische Dienste Aargau AG

Comment number	Chapter	Comment	Suggested change	Response
General Comments				

Comment number	Chapter	Comment	Suggested change	Response
38	General	The main point of criticism is the lack of evidence for the effectiveness of the interventions examined for different degrees of MDD severity. Previous research has clearly shown that whereas CBT is superior to ADM in slight and moderate MDD cases, ADM or combination therapy are preferred therapy options in severe MDD. The only comment about the effectiveness of ADM in severe cases in on page 14 ("The pooled analyses of relapse in ADM compared to placebo showed a lower risk of relapse with ADM, also in patients with severe MDD"). In the "Conclusion" no intervention is clearly recommended: "Within the available data, we found no evidence of a difference between ADM, CBT and their combination with regards to the primary clinical effectiveness outcomes". This result may reflect the lack of severity-specific evaluation of the data. Nevertheless, we support the finding about "the favourability of CBT over ADM and their combination over monotherapy"		Thank you for your valuable feedback. We agree that exploring the differential effects of ADM and CBT across the different MDD severities is certainly important. However, as you said, the low number of studies unfortunately did not allow for such an evaluation.
39	Results-Assessment of Clinical Effectiveness and Safety	We support the statement "that ADM alone was not optimal from the perspective of acceptability and tolerability".	-	No response needed.

Comment number	Chapter	Comment	Suggested change	Response
40	Age-related effects and tolerability	Unfortunately the findings do not provide any evidence for effectiveness or tolerability of either interventions or combination therapy in elderly patients.		No response needed.

Full name	Birgit Watzke, Markus Wolf
Function	Full Professor, Senior Researcher
Organization	University of Zurich

Comment number	Chapter	Comment	Suggested change	Response
General Comments				
41	1; 2	Reporting and writing style in the Executive Summary and Introduction are sometimes too vague or imprecise and do not fully reflect the literature or empirical evidence from the current analyses. Specifically, there are two instances that should be reconsidered: On p. 14 it reads “The pooled analysis of relapse in ADM compared to placebo showed a lower risk of relapse with ADM, also	Claims made in the Executive Summary should be carefully evaluated because most readers will not read the full report but rather take relevant information from this summary. In addition, it would be helpful to harmonise writing style and content within and between the Sections A, B and C of the report, e.g. in terms of when and how exact numbers for important outcomes are reported etc.	Many thanks for your detailed comments and valuable feedback. We have revised the wording for the two points mentioned: 1. *“The pooled analysis of relapse in ADM compared to placebo showed a lower risk of relapse with ADM during the continuation phase.”* (Executive Summary, section A of results, page 15). 2. *“Current guidelines recommend psychotherapy as initial treatment in those with mild to moderate depression and the combination of antidepressants and*

Comment number	Chapter	Comment	Suggested change	Response
		in patients with severe MDD". A closer look at the two studies that went into this subanalysis reveals that the evidence base with regard to severe MDD is very weak, partly misleading and should not be presented in such a prominent way. In addition, the statement on p. 20 needs further clarification from our perspective, claiming that "Most guidelines recommend their [ADM] use as first line therapy, either alone or in combination with psychotherapy, particularly in those with moderate to severe symptoms (16–21)." From our perspective, based on the evidence most current guidelines do not recommend ADM in all of these cases, but rather recommend psychotherapy for mild to moderate depression, and a combination of CBT and ADM in severe MDD.	Especially, it could be made explicit which aspects of the sections A, B, C were similar and in which aspects the three sections were somewhat independent, i.e., for which of the three parts a separate literature search was conducted and why.	*psychotherapy in those with moderate to severe symptoms."* (section 1, page 22). We have also revised the Executive Summary and attempted to harmonize the writing style and presentation of the results.
42	1; 5.2.1	One of our main concerns refers to the scope and type of research included in the review, and the potential risk that substantial body of research was not addressed though highly relevant with regard to the overall aims of the HTA. The aims are described on p. 10: "In this Health Technology Assessment	There are a number of CBT trials that utilize a smaller number of sessions or less-than-13-week interventions (e.g., 8-10 weeks) but which specifically aim at continuation and/or relapse prevention in MDD (e.g., Bockting et al, 2005; 2009; 2018; Fava et al., 1998; 2004; Jarrett	Thank you for raising this valid point. We certainly agree with you that an important consideration is whether acute phase CBT treatment has enduring effects on relapse prevention after discontinuation and how that compares to maintained antidepressant treatment. However, our aim was focused on whether maintained treatment of either CBT or ADM was more efficacious, rather than the enduring effect of either. We apologize if this was not sufficiently clear in the aims. We are now explicitly mentioning this in the report (for example, *"In this Health Technology*

Comment number	Chapter	Comment	Suggested change	Response
		(HTA), we aimed at evaluating the clinical and health economic effectiveness, safety, and benefit-harm balance of ADM and CBT interventions, alone or in combination, beyond the acute management phase (i.e., >12 weeks)." We understand that the aim was to assess enduring effects of ADM and CBT on relapse and other clinically valid outcomes beyond the usual 12-week acute phase, which is a highly relevant question. What remains unclear to us from reading the introduction and rationale of this HTA, was the reason for excluding studies that investigated this question but tested CBT interventions that were specifically designed for this purpose but were shorter than 13 weeks. From the flowchart (Fig. 3) it becomes clear that a substantial number of studies were excluded because their intervention was shorter than 13 weeks; in this context we are wondering what the rationale was to include interventions with >12 weeks duration only. Of course, most	et al., 2001; Ma & Teasdale, 2004; Paykel et al., 1999; 2005; Segal et al., 2010; Shallcross et al., 2015; Stangier et al., 2013). Even in case the HTA had a slightly different focus - which we think should then be made more explicit - this substantial body of research should be taken into account when discussing the long-term effects of CBT on relapse. Interestingly, shorter duration trials with a shorter treatment duration seem to be included in the cost-effectiveness analyses. What were the reasons for the different procedures? We recommend that the reasons for these different approaches should be made more explicit in the introduction, methods and discussion because otherwise the data base varies considerably between the different parts of the HTA. In case these aforementioned trials were excluded for a reason other than treatment duration, this should also be made more explicit in the text. Finally, it	*Assessment (HTA), we aimed at evaluating the clinical efficacy and safety, benefit-harm balance and health economic characteristics of ADM and CBT interventions, alone or in combination, in patients with MDD and receiving these interventions beyond the acute management phase (i.e., >12 weeks).").* While it the enduring effect of CBT is outside the scope of this report, we briefly mention it in the section 5.3.2: *"It is worth noting that some studies have shown that acute phase CBT may have an enduring effect even after the discontinuation of acute phase treatment. This is important as patients are usually maintained on ADM for prevention of relapse and recurrence. A comparison of the long-lasting effects of both CBT and ADM (in their presence and absence) would allow for better decision making by patients and clinicians in choosing the initial treatment. However, we did not address this aspect in this review as we only included studies were participants were maintained on either treatment."* We have also added a list of all excluded studies with reasons (Appendix 11.1.3, page 34). Regarding the discrepancy in the study inclusion criteria in the clinical and cost effectiveness sections in terms of CBT duration, this is related to the availability of data as most studies identified in the health economic assessment were of shorter duration. Sections 7.1.1 and 7.3 reported the main differences between the economic and clinical reviews and we

Comment number	Chapter	Comment	Suggested change	Response
		psychotherapies last longer, but quite a substantial number of trials that have dealt with the topic of this review, i.e. the preventive effects of (shorter-than 13 sessions) CBT on relapse in patients with a diagnosed depression, but they seem to have been excluded. Given that there is no empirically agreed-upon standard number of sessions or weeks for psychotherapy to be effective, we are wondering whether this restriction is creating an artifact neglecting important evidence from the relapse prevention literature.	would be good to know which studies were excluded for which reason, so we recommend to include a table in the appendix listing all excluded studies by reason for non-eligibility (similar to the one presented for the cost-effectiveness analyses, see 11.3.2). Finally, we recommend that the authors point to this important body of research when discussing future clinical directions for the long-term management of MDD.	have added a statement regarding the duration of the trials: *"It is important to note that there were several methodological differences between the economic systematic review and the clinical systematic review performed in this HTA….. Unlike the clinical systematic review which only included CBT interventions which lasted for at least 12 weeks, we included two cost-effectiveness analyses of CBT that reported a treatment duration potentially below 12 weeks (the authors reported 6-8 weekly CBT lessons/modules but did not specify whether the lessons/modules were conducted consecutively in the shortest time possible, or whether they were distributed over a period of time exceeding 12 weeks).In general, it should be emphasized that in many cases the real duration of CBT was not specified in months, but in number of sessions or modules. The exact distribution of sessions/modules over time was often not reported."* We have further added the following in the Executive Summary to outline the differences in the clinical and economic reviews: *"It is important to note that were some methodological differences between the clinical and economic reviews of this HTA including the inclusion of patients with substantial depressive symptomology in the economic review, in contrast to the clinical review which focused on patients diagnosed with MDD. While the clinical part included all international studies, the economic part excluded East Asian countries as they are likely to substantially differ to Switzerland in terms of settings, costs, and perception of quality of life. Furthermore, unlike the*

Comment number	Chapter	Comment	Suggested change	Response
				clinical systematic review which only included CBT interventions which lasted for more than 12 weeks, we included two cost-effectiveness analyses of CBT that reported a treatment duration potentially below 12 weeks (the authors reported 6-8 weekly CBT lessons/modules but did not specify whether the lessons/modules were conducted consecutively in the shortest time possible, or whether they were distributed over a period of time exceeding 12 weeks). In general, it should be emphasized that in many cases the real duration of CBT was not specified in months, but in number of sessions or modules. The exact distribution of sessions/modules over time was often not reported."
43	4.2, 7	The included types of CBT are quite heterogeneous, not well introduced, and not consistently addressed in the different parts of the HTA. For instance, it is stated that "blended" CBT was included, but it is unclear exactly what this means (e.g., were unguided internet-interventions excluded)? Overall, relatively few information is given about the various forms of CBT that were finally included. Finally, there are some contradictory statements, e.g., on p. 101 it reads that CBT was administered for "at least 4 weeks", while it reads "at least 6 weeks" on p. 102.	Maybe make more explicit, and describe in more detail the interventions included in the different chapters of the review (type of CBT, session numbers, form of delivery, conducted by licensed therapists, bona fide…). This would help the reader to understand what form of "CBT" was actually analysed in the different parts of the HTA, and to assess how this translates to real-world psychotherapy. For instance, it is questionable if 4-session internet-based treatment should be considered "CBT" to treat patients with MDD as	Information regarding the details of the CBT interventions in the clinical studies has been provided (see appendix 11.1.3) and the impact of including heterogenous interventions in our analyses and implications for the robustness of the estimates have been additionally discussed (section 5.3.2.1) – please see comment 2 or 20. The differences in the methods of the clinical reviews are outlined in sections 7.1.1 and 7.3 and are now also briefly reported on the Executive Summary (please see comment 42). We apologize regarding the contradictory statements on CBT duration. We have corrected it on page 102 to "at least four weeks".

Comment number	Chapter	Comment	Suggested change	Response
			suggested by the title and aims of this HTA. In the executive summary and overall conclusions it should be highlighted that quite different inclusion criteria of CBT were used in the different parts of the HTA.	
44	1; 5.1; 6.1; 7.1; 8.1	In the Executive Summary, Sections A, B, C are somewhat heterogeneous in the way procedures and details are reported. There is also heterogeneity in the overall structure of the HTA and the way of reporting. Despite similar subheadings the type and amount of information given in chapters 5, 6, 7 and 8 varies which makes it difficult for the reader to follow, and specifically to spot substantial methodological differences between these major parts. Sometimes it is unclear if certain methods/procedures were the same, or whether and why they differed between the different chapters (5.1-8.1). In the Executive Summary, it is stated that this HTA mostly focusses on the level of treatment	For easier comprehension, all paragraphs should be harmonised with regard to the type, amount, as well as level of elaboration, of information and should follow the same reporting structure for methods description. It would be good if it wouzld be made more explicit what was similar and what differed between effectiveness, BHA, HEA, and budget impact analyses. In addition, the transitions between chapters could be improved.	See response to comments 41 and 42.

Comment number	Chapter	Comment	Suggested change	Response
		modalities (ADM vs. CBT) which is also reflected in most parts of the HTA including the Executive Summary. However, the cost-effectiveness paragraph discusses distinct antidepressants quite extensively. Similarly, while type of sponsorship is explicitly mentioned in the cost-effectiveness para no comment on this is made in the effectiveness part.		
Specific Comments				
45	2	"While many of the commonly experienced adverse effects are often transient and self-resolving (35), little is known about their persistence or development with long-term use"	To us this reads a bit contradictory: If there is lacking knowledge about long-term AE, how can be stated that they are "often transient" and "self-resolving". Maybe reformulate.	The sentence has been rephrased to: *"While some of the commonly experienced adverse effects, such as nausea and headache with ADMs, are often transient and self-resolving in the short term, little is known about whether long-term use of ADM or CBT is associated with the persistence or development of other adverse effects."*
46	4.4.1	It is stated that relapse is used "as defined by each of the studies". This could result in rather heterogeneous outcomes.	Maybe briefly describe the way and range how relapse was defined (e.g., compared to the standard definition, see Fig. 4) in the included studies to enable the reader to understand the differences and impact that these definitions might have on the outcomes.	The definitions of relapse for every study are included in Table 2 in the main results. Additionally, we have now added the definitions used for response and remission (Table 12 and Table 13).
47	4.4.2; 6.1.3	It is stated that "acceptability is measured as the proportion of participants who left the trial for any reason prior to the end of the study" Does this include discontinuation of treatment/treatment dropout as opposed to study dropout, or loss-	Maybe more explicitly define what is meant by acceptability, tolerability, e.g., differentiate between study dropout, treatment dropout/discontinuation.	Acceptability included study drop-outs for any reasons and tolerability included study drop-outs due to side effects.

Comment number	Chapter	Comment	Suggested change	Response
		to-follow-up? Some patients prematurely stop the intervention without dropping out of the trial, i.e. they continue to provide outcome data. A similar question is related with the term tolerability. On p. 88, it further reads “we also considered all-cause dropout (i.e., treatment discontinuation for any reason)”. Does this include study dropout, or loss to follow-up?		
48	4.4.2	It is written that “We chose the primary clinical and safety outcomes in consultation with an external clinical expert.”	Because this is a crucial decision, we suggest to describe this process in more detail, and probably disclose the individual(s) involved in this process to increase transparency (or/and disclose their professional background and expertise). Specifically, was this decision based on a consensus process, or was it informed by the consultation meeting, or the reviewer comments from earlier project stages?	The clinical efficacy, safety, and cost effectiveness outcomes were agreed on in consultation with the experts and stakeholders during the scoping process. The primary clinical and safety outcomes were chosen based on what is commonly assessed in trials evaluating mid-long term treatment of MDD (for example, relapse and recurrence) and what could be possibly relevant to both clinicians and patients (for example, quality of life) when deciding on the treatment options. This decision was based on the literature and on a discussion with Prof. Dr. Ulrich Schnyder who is a specialist in psychiatry and psychotherapy. We have further extended our description of the process of selection of outcomes in section 4.4.2. and all experts which have participated in the process are now acknowledged in page 4.
49	5.1	It is stated that “We were not able to assess publication bias using funnel plots due to the low number of studies”	Given that trials with negative findings are often not published, and certain quality measures such as pre-registration and publication of protocols were not very common 10ys ago (at least in the field of psychotherapy research), this aspect should be	The following has been added to the discussion section 5.3.2.1: *“The efficacy of ADMs has been reported to be overestimated because of publication bias and selective outcome reporting. This may occur due to several factors including reluctance of authors or sponsors to report and journals to publish null*

Comment number	Chapter	Comment	Suggested change	Response
			briefly addressed in the discussion. If possible, maybe provide a brief evaluation whether or not the authors think if publication bias was likely a problem in the two treatment arms (ADM, CBT).	*findings, especially in the early 2000s before trial registration was enforced. Similarly for psychotherapy trials, there is a growing body of evidence that the clinical benefit of psychotherapies reported in meta-analyses may also be overestimated because of publication bias. However, we were unable to assess the extent of publication bias and selective outcome reporting due to the small number of trials included and both biases remain likely for both ADM and CBT and the efficacy of both may have been overestimated."*
50	5.2.1	Of 15 identified eligible reviews, two high-quality reviews were selected to identify individual RCTs.	Once, all the reviews were retrieved, why did the authors not screen all reviews, including the most recent ones?	We conducted our search and review and study selection according to what was specified in the scoping protocol. However, we agree that the screening of all reviews could have potentially led to the identification of additional studies. The following has been added to section 5.3.2 : *"We conducted a comprehensive search process in multiple databases and used two of the largest reviews on antidepressants and cognitive behavioural therapy to identify relevant studies. We used the same search strategies as used by the authors of the two reviews to ensure consistency, accepting that some of the search terms were incomplete (e.g., LU AA 21004 was not contained as a vortioxetine search term). While we are confident that our search strategy captured relevant studies well, we cannot exclude that some may have remained unidentified."*
51	5.2.1	According to the flowchart n=44 studies were finally included. But as far as we understood, the sample for the different outcomes differed from that number and only 9 studies	Maybe include the numbers that were finally available for each of the outcomes (e.g., n=9 for relapse).	The numbers of studies included for each outcome have been included in the updated figure.

Comment number	Chapter	Comment	Suggested change	Response
		were used to assess the primary outcome of this HTA.		
52	5.2.4.2	In the last para, on p. 73, it reads "Pooled analysis of these two studies resulted in an RR of 0.66 (95% CI= 0.28 to 1.58) with a tendency to favour ADM..."	A look at Fig. 24 suggests that this sentence should read "Pooled analysis of these two studies resulted in an RR of 0.66 (95% CI= 0.28 to 1.58) with a tendency to favour ADM plus CBT ..."	Thank you for pointing this out. The sentence has been corrected.
53	5.3.2.2	Risk of bias was considerable in the included data set.	Maybe address sources of bias in more detail in the text, and also consider to report RoB separately for ADM and CBT studies.	Further details on the risk of bias in the included studies have been added to the discussion section 5.3.1 *"The studies additionally suffered from several methodological shortcomings with most studies being at high risk of bias in several domains. This included possible selective outcome reporting and lack of blinding and allocation concealment. In the six CBT trials where double blinding was not possible, only two trials reported blinding of outcome assessors. Furthermore, because of this high risk of bias, our GRADE assessment showed that the quality of evidence of studies which included a comparison of ADM and CBT needs to be regarded as low. Insufficient blinding in these trials may have biased the estimates in the direction of CBT."*
54	5.3; 9	We are wondering whether the conditions for unbiased network meta-analyses are met in this dataset given the small number of quite heterogeneous studies. Specifically, transitivity could be violated due to the fact that the trial designs differed on many factors (e.g., the way control groups, specifically TAU, were defined).	The validity of NMA results should be critically reflected in the limitations.	We agree that the transitivity of the network may likely be affected by the heterogeneity of studies, particularly in terms of designs (lack of blinding in studies evaluating CBT) and control groups used. While we tried to ensure that the transitivity assumption was not violated to the extent possible given the available data (i.e., by excluding studies assessing the effect of interventions in patients with specific medical comorbidities as well as performing analyses for studies that were similar in design), we

Comment number	Chapter	Comment	Suggested change	Response
				cannot exclude that other factors such as different components of the comparator groups (e.g., different treatments may have been included in the TAU groups across different studies) differed among the studies, affecting the validity of the network estimates. Due to this uncertainty regarding the validity of the network results, we considered the primary analyses to be the pairwise meta-analyses and all findings were interpreted based on the results of these analyses. A discussion of this and other limitations has been added to the revised report (please see comments 1 and 20).
55	5.3; 9	Given the tremendous economic impact (in Switzerland) of direct costs related with hospitalizations due to MDD, as well as steadily growing indirect costs (loss of productivity due to long sick leaves), one of the conclusions could be that the prevention of these consequences should be a top priority in future depression treatments and research.	The overall number of affected individuals and intense inpatient treatment are the main cost drivers. In the discussion, maybe mention hospitalization as future key outcome criterion (as well as additional criteria that better address sustained, real-world outcomes more comprehensively than single-event outcomes such as relapse, e.g. well-being, days on sick leave, early retirement). In addition, taking the evidence from this HTA into consideration, the authors might give a brief resumee and outlook about which treatment strategies they consider most likely to be effective in Switzerland to tackle these challenges on a population level, i.e., to increase the population-based impact of	We agree that hospitalizations are a major cost driver and need to be considered as an important outcome in evaluating treatments of MDD. We have added a statement according to your suggestion: *"This finding highlights that hospitalisation due to MDD is an important clinical outcome that, along with other consequences for occupational activity and well-being, should ideally be captured by clinical trials and observational studies to allow for a more comprehensive assessment of MDD treatments."* This document represents the scientific assessment part of this HTA. As per standard SMB processes, it is not within the scope of this report to issue recommendations, which is the task of the appraisal committee.

Comment number	Chapter	Comment	Suggested change	Response
			available depression treatments, to prevent depression onset, and to reduce the number of hospitalizations.	
56	6.1.4	In some parts of the HTA it is challenging to capture which studies went into which result. For instance, in Table 13 it would be good to know the studies behind the numbers.	For better transparency we suggest to include the number of studies and references that reflect the data given in each of the cells of the tables (e.g., Table 13). We would also recommend to cite references in the text where results are presented based on a certain group of studies.	Where possible, we have added references to the text and tables.
57	7	In this chapter, the included CBT interventions appeared to be very heterogeneous, many being internet-delivered, and some located at the very low-dosage end of the continuum. Analyses based on these interventions need to be reflected in terms of generalizability to real world psychotherapy. Similarly, the cost-effectiveness analyses included patients with “elevated symptoms” which are not included in other parts of the HTA (this decision might coincide with the high number of internet interventions in this part of the HTA). So it seems that different parts of the HTA are addressing different populations which does not become fully clear from the way this part is embedded in the HTA	Maybe treat findings from bona-fide CBT and low-intensity/ self-help/iCBT separately, and more explicitly address the fact that quite different populations were looked at in the different parts of the HTA. Maybe critically discuss the specific selection of studies and the limited real-world implications of this selection, specifically the high proportion of brief internet-interventions that are often more of an online variant of a self-help book, rather than full-blown psychotherapy.	We pointed this out in the limitations section (chapter 7.3.): “One main limitation affecting the interpretability of the available cost-effectiveness results concerns the high heterogeneity across the identified studies, which included different populations (from patients with symptoms related to depression to patients with severe MDD), different interventions and comparators, different time horizons, as well as different types of costs. Especially for CBT interventions there were significant differences across the identified studies.” Moreover, in the conclusion (chapter 7.4.) we clearly state that “The high heterogeneity in the results of the CBT studies suggest that the cost-effectiveness of this is intervention may depend on how CBT is provided (e.g. number of sessions, treatment duration, setting).”

Comment number	Chapter	Comment	Suggested change	Response
		(Executive Summary, Introduction, etc).		

Reviewer Comments

Full name	**Stefan Felder**
Function	Professor in Health Economics
Organization	Department of Health Economics, University of Basel

Comment number	Chapter	Comment	Suggested change	Response
58	Title		Antidepressants and cognitive behavioural therapy interventions of major depressive disorder	Thank you for your comments. Change accepted.
59		As I explained in my comments, “or” includes both options. This is basic set theory learned in primary education. The set union includes the intersecting set. I know that there are a few people that write “and/or” but that makes the case not better – quite to the opposite.		Change accepted.
60		You either delete that or inform the reader that also the clinical studies were sponsored by pharmaceutical companies. It is not sufficient that the information of sponsorship of clinical trials is given later in the text.		Sponsorship status of trials has been added to the clinical part of the executive summary.
61		If I’m not mistaken, the threshold for cost-effectiveness is not provided. But this is neede if you assess a treatment as cost-effecitve.		The sentence where the treatment is assessed as uncertain in its cost-effectiveness has been deleted, as the thresholds for cost-effectiveness vary between the different countries the analyses were performed for. Also, the references to Annemans, and Ramsberg have been added.