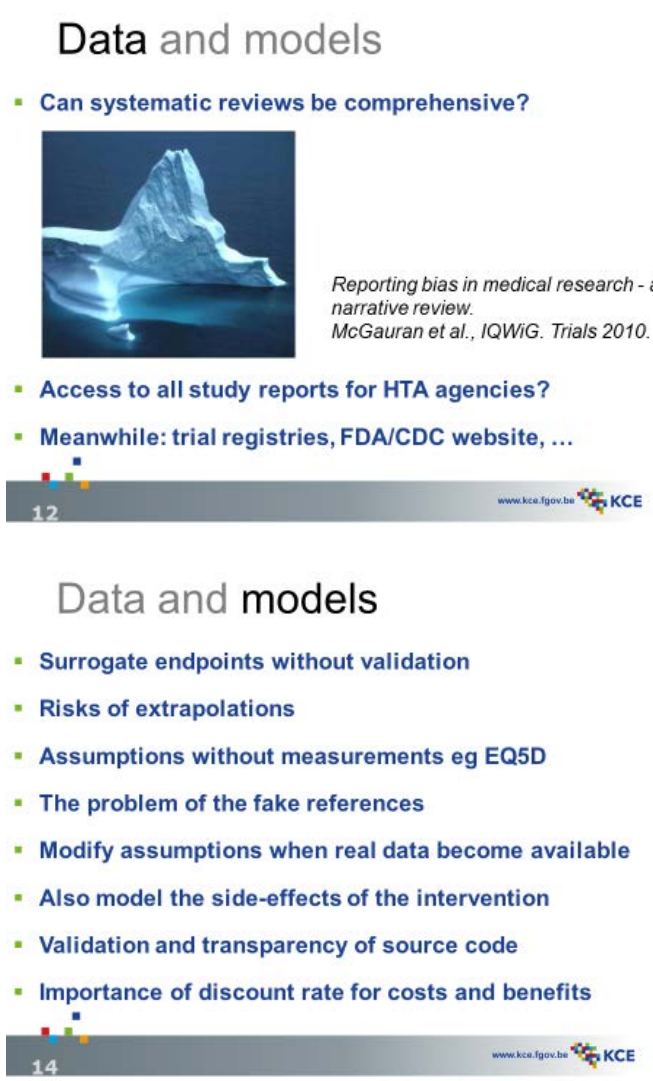
JA2- WP7- SG 3 – **SAG & Public** consultation on the methodological guideline "Methods for health economic evaluations" - Draft Guideline

Number of com-ment	Page	Line	Comment	Character of comment "major"[1] "minor"[2] "linguistic"[3]	Author/Draft group reply
1	GEN		Thanks for sending the review. We have looked through these and don't find anything that requires comment. Thanks for the effort and sharing.		Thank you for your review!
2	GEN		As a general comment to the whole document we are concerned with the fact that the guideline does not link common agreement across Europe with best practice. We understand that the guidelines merely collated information about individual countries behaviour and practices and in doing so considers areas of common agreement/consensus between countries (with respect to the conduct of economic evaluation). However, consensus does not, in our view, necessarily imply that the best methodology is followed. Thus, in the spirit of providing guideline over and above the consensus, it would have been perhaps helpful to link *current* practice to establishing *best* practice. We acknowledge the fact that, potentially, countries may be justified in following a specific set of criteria guiding their own behaviour, but this is perhaps beyond the scope of this document.	☒ major ☐ minor ☐ linguistic	Thank you for this advice. We do agree that what you ask for would have been good, but it is, as you say, beyond the scope of this document. The purpose of this document is to guide assessment teams within EUnetHTA on how economic evaluations can be conducted to be as useful as possible to decision makers in individual countries.

[1] "major" indicates that a comment points to a highly relevant aspect and that the author / the draft group is expected to give a thorough answer
[2] "minor" means that a given comment does not necessarily have to be answered in a detailed manner
[3] "linguistic" labels problems with grammar, wording or comprehensibility

3	GEN		EFPIA would like to provide some general comments to the public consultation on the methodological guidelines "economic evaluations". EFPIA would also like to comment to the process in relation to the governance and timing of the consultation. Unfortunately, it appears that the SAG WP 7 stakeholders did not have a proper opportunity to review the methodological guidelines before it was sent for public consultation.		Thank you for your comments! As laid down in the WP7 workplan for Joint Action 2, the development process for methodological guidelines in EUnetHTA has been revised. In case of unforeseen difficulties with this new, more interactive elaboration model, two consecutive 14 months periods of guideline elaboration should provide an opportunity for interim corrections. The ambitious, 14 month development process in JA2 does not allow the conduct of a separate phase of SAG consultation on the draft guidelines before the public consultation phase. It has to be run in parallel. The only current „privilege" for the SAG is an extended consultation period by two weeks. The future model of methodological guideline development in EUnetHTA will be based on the experiences in JA2. It will probably be more flexible in regard to the time available for guideline elaborations, and it will contain a separate SAG consultation period completely finalized before the start of the public consultation." (Answered by SG3 coordinator)
4	GEN		EFPIA is of the opinion that the proposed guidelines aim at the lowest common denominator of all national guidelines and may as such not be very helpful, but serves as a systematic review of what exist in the field. Differences between national guidelines are mentioned to illustrate the breadth of acceptance of these		Thank you for this advice. We do agree that what you ask for would have been good, but it is beyond the scope of this document.

			recommendations, for example regarding surrogate outcomes, preference elicitation and mapping. Interestingly, a wide range of modelling techniques and tools are acknowledged, however accompanied by a recommendation of transparency and of providing the technical model to the agencies. It would have more impact if EUnetHTA took a principled stance on some of the methodological issues with an aim of producing economic evaluations with maximal transferability and relevance across countries in order to reduce duplication of work in the different member states.		
5	GEN		It is not clear to me whether you want to summarize only the theory from the various HTA agencies or also want to provide the reader with common caveats seen in modelling. I feel this is missing. I have listed for your information the two slides I present to industry and others who may be interested. I have examples for every line illustrating the importance for the conclusion of the model.		Thank you for your review! It was not in the purpose to provide the reader with common caveats seen in modelling. We agree that it would have been good, but it is beyond the scope of this document.

					
6	6	127	Purchasing power parity – not plural	☐ major	Thank you, we have revised this accordingly.

				☐ minor ☒ linguistic	
7	8	153-154	– Statements about the objective of the individual economic evaluation should be added in the EUnetHTA guideline as HTA agencies have different remits across Europe.	☒ major ☐ minor ☐ linguistic	We do present the purpose of conducting economic evaluations in the EUnetHTA countries in Table A3.
8	8	155-157	As other guidelines for other interventions than pharmaceuticals are mentioned here (i.e. medical devices), it seems important to state as well that some specific HE guidelines exist for vaccines, which have several specifities due to their infectious status: - WHO guide for standardization of economic evaluations of immunization programmes: http://whqlibdoc.who.int/hq/2008/WHO_IVB_08.14_eng.pdf - JCVI code of practice in the UK: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/224864/JCVI_Code_of_Practice_revision_2013_-_final.pdf - On-going work in Germany on vaccines specific HE guidelines: http://www.rki.de/DE/Content/Infekt/Impfen/Forschungsprojekte/STEErING-Projekt/STEErING-Projekt_node.html	☒ major ☐ minor ☐ linguistic	We fully agree that there are specific issues in guidelines on vaccines, eg, dynamic models etc. However, there are very few guidelines for health economic evaluations solely pertaining to vaccines and none of the contact persons of the EUnetHTA members have provided information about specific guidelines for vaccines, even though we have asked for guidelines for all different types of health technologies. It is also our understanding that the JCVI code of practice cannot be understood as a guideline concerning health economic evaluations and that it refers to other NICE guidelines on these issues. However, we have based on your comment decided to, as an example, include additional information from one of the guidelines (France) concerning discounting in relation to the evaluation of public health programmes such as vaccines. The German work for the Robert Koch Institute and the German Standing Committee on Vaccination has just recently started. The WHO guidelines will be mentioned under the heading "Other related documents".

9	8	158	– *Replace:* "standpoints" with "viewpoints".	☐ major ☐ minor ☒ linguistic	We have revised this accordingly.
10	8	168	The document – already in the summary – correctly notes that there are several aspects where commonalities do not exist, yet the document continues to act as if this does not present an issue. In fact, it does undermine the general purpose. It would seem better to try sharing inputs for the evaluation than outputs.	☒ major ☐ minor ☐ linguistic	Thank you for this comment. However, we do not completely agree with you. Even when commonalities cannot be found, it may still be useful to present the different methods that are recommended, and try to find a way of presenting the results in a way that makes it more useful to different users (e.g. through sensitivity analysis).
11	8	175	By not addressing differences, a common view on conducting health economic evaluations is not actually the aim of this document.	☒ major ☐ minor ☐ linguistic	It has not been the purpose of this document to find a complete common view on conducting health economic evaluations. The two main aims have been: "To increase the knowledge about similarities and differences between guidelines for health economic evaluations, used in European countries" and: „To develop a common framework for the methodology of economic evaluations for EUnetHTA based on the identified commonalities."
12	9	178	– *Replace:* "concern" with "represent".	☐ major ☐ minor ☒ linguistic	We have revised this accordingly.
13	9	182 REC	**Evidence of clinical effects** It should be recognized that different methodologies for studying		The recommendations are based on the commonalities found among the guidelines. We have tried to clarify that other study designs may

		2	treatment effects (randomized controlled trials, pragmatic trials, observational designs etc.) have different and complementary strengths and weaknesses. The quality of a study should be assessed based on its ability to inform decisions about the appropriate use of the new technology, not according to any fixed hierarchy of evidence. HTA is about the value of technologies in routine care, thus the value of methodologies that examine the outcomes with technologies used under the conditions of routine care needs to be emphasized.		contribute with complementary information.
14	9	182 REC 3	**Time horizon** The recommendation that the time horizon should be "sufficiently long" may not be very helpful in practice. The choice of time horizon is intrinsically linked to which downstream events are included in the analysis. The decision on which events to include needs to consider available evidence. E.g. if events are included for which data on treatment effects is weak, this will inflate uncertainty without improving cost-effectiveness estimates.		Once again we agree with you, but despite the vagueness "sufficiently long" is the word that is most often used in the regional guidelines.
15	9	182 REC 4	**Modelling** The guidelines state that modelling should be conducted "when methodologically appropriate", but do not state under which conditions modelling are appropriate. It would be better to recognize that modelling is a practical and desirable necessity in the conduct of economic evaluation.		We agree with your comment and have removed the word "methodologically" from this sentence. In this short recommendation we do not want to go into the question of when modelling is assumed to be "appropriate", but examples are given in section 2.3.4.

			Estimating the cost-effectiveness of a new technology involves combining data from several sources and making predictions of future events. Such analysis requires assumptions about how short-term clinical effects are translated into economic and humanistic outcomes over the long term. A model is an explicit formulation of assumptions made to estimate cost-effectiveness from available data. It is desirable that assumptions are clearly and transparently stated so that calculations can be independently verified and the validity of assumptions can be tested. A model provides such a framework, and is to be preferred over incomplete or non-transparent analysis, which would be the alternative to modelling.		
16	9	182 REC 5	**Perspective of economic evaluation** To ensure transferability of results, economic evaluations should include all effects and all costs, irrespective of payer (societal perspective). Health systems are organized differently, and which resources are included in a "health care payer perspective" will vary from country to country (e.g. the costs for long-term care may or may not be included). In several countries the inclusion of production costs is recommended, thus these costs should be included. All costs should be disaggregated in prices and quantities and presented separately. If a treatment affects survival, then resource use during the increased life expectancy should be included in the economic evaluation. The suggestion to distinguish between "related" and		You have many good points, but our aim was to compare the recommendations in the guidelines, and investigate if there is a common view, or a way that the results could be presented to make it more transferable between EUnetHTA partners.

			"unrelated" resource use is misleading; if the change in resource use is caused by the treatment then it is related. There is no reason why changes in resource use due to mortality effects should be included differently than changes in resource use due to morbidity effects. In cost-utility analysis, QALY gains due to mortality effects are included in the same way as QALYs gained due to morbidity effects (health gains).		
17	9	182	„To enhance the usability of the economic evaluations, it is recommended that results be presented in terms of both cost-effectiveness analysis (CEA) and a cost-utility analysis (CUA)." This is the first of many examples where the guidance is to do everything because there is no consensus. This does not make for unified decisions but instead opens up for submitting multiple types of analyses to one agency solely for the purpose of the analyses being shared, even if the agency only requests one type of analysis themselves.	☒ major ☐ minor ☐ linguistic	For analyses performed within EUnetHTA to be of value to as many EUnetHTA members as possible, this recommendation would be of help. Many countries recommend the use of QALY, but other countries do not recommend it, so there can be no type of analysis that suits everyone for the moment.
18	9	182	The recommendation summary table does not address the most important international differences; differences in health care setup, cohorts and local costs and clinical settings. The analyses will therefore not reflect the decision needed to be made.	☒ major ☐ minor ☐ linguistic	The differences you mention cannot be said to be methodological and are therefore not within the scope of this guidelines. However, we agree that it is important to point out that these differences also exist and we have tried to clarify this in the introduction and the conclusion.
19	9	182	Recommendation 1: A more specific guidance as to what type of health economic analysis should be used would be helpful	☒ major ☐ minor ☐ linguistic	For analyses conducted within EUnetHTA to be of value to as many EUnetHTA members as possible, we would argue that the existing recommendation about type of analysis would be of help. Many countries recommend the use of QALY, but other countries do not recommend it, so there can be no type of analysis that suits

					everyone for the moment.
20	9	182	Given that it is often not possible to compare against all products within a disease area, guidance on the selection of comparators for the health economic analysis would be helpful.	☒ major ☐ minor ☐ linguistic	Selection of comparators is covered in section 2.2.1. This was, however, not considered one of the 10 main recommendations.
21	9	182	Use of the term: "quality of studies" would be better represented by focussing on the "critical appraisal of studies risk of bias" in line with the Cochrane Handbook 2011. The handbook distinguishes clearly between "risk of bias" and "methodological study quality" and recommends a focus on *risk of bias* (most importantly, high methodological quality does not remove potential for risk of bias and their is a potential ambiguity between quality of reporting and quality of the underlying research). Julian PT Higgins, S. G. (2011). Cochrane Handbook for Systematic Reviews of Interventions. The Cochrane Collaboration.	☐ major ☒ minor ☐ linguistic	We agree that risk of bias may be a better term for critical appraisal of studies but have decided to stay with the term quality of studies since it was considered a broader term that also includes bias minimisation. In addition, the guidelines we reviewed mention more often the "quality of the studies" than "evaluation of the risk of bias".
22	10	182	We suggest rephrasing the second sentence. The sentence implies that because most countries use a discount rate of 3 to 5 percent, the recommendation is that costs and effects should be discounted in the base-case. There is a clear rationale for the theoretical basis in the health economics literature which should be the rationale for recommending discounting here and the appropriate steps of obtaining such discount rates.	☒ major ☐ minor ☐ linguistic	The purpose of the guideline is to find methods that make the economic assessments useful for as many EUnetHTA members as possible. We have removed the word thus, since we considered this word to be unnecessary.
23	10	182	Recommendation 8: It would be helpful to state what specific discount rate should be used for costs and for effects. The	☒ major	As different countries recommend different rates, we are not in a position of recommending

			current range listed is not a sufficient guidance and it is not stated whether the discount rate should be identical for costs and effects.	☐ minor ☐ linguistic	one rate. This could perhaps be the focus of future EUnetHTA projects.
24	9	182	"… may be conducted…" Which decision makers request how often health economic analyses for what type of technologies across Europe is a valuable question that should be answered (potentially by the strategic HTA Network) to give priority to useful and fit-for-purpose future developments in this field. In times of economic challenges, EUnetHTA and upcoming Joint Actions should focus its' resources on topics where there is a real request from end users of EUnetHTA work products. A business case for health economics and HTA in general should depend on the actual use for decision making (ref. Grant Agreement of Joint Action 2). Usage needs to be tracked. Stakeholders need to be involved to input to the appropriateness and data basis of such assessments across Europe.	☒ major ☐ minor ☐ linguistic	We will forward this comment to the lead partners of the work package.
25	9	182	The conclusion that a systematic review should be the basis is the clinical effectiveness review does not comply with the current practice as displayed in table A13: 11 of 21 guidelines accept RCTs as the basis without a systematic review. It is not surprising that the producers of systematic reviews recommend their use (this is only a statement derived from a self-assessment), but is there a real business case for HTA and for	☒ major ☐ minor ☐ linguistic	Our recommendation is based on this line „the majority of the countries with guidelines state that they prefer systematic reviews and meta-analyses". Furthermore, none is against it. Even though many do accept RCTs without a systematic review, this does not mean that they recommend it. We acknowledge that there may be specific problems related to the assessment of medical devices. However, our

			which technologies? To what extent is HTA really used for the different types of technologies across Europe and to what extent do decisions follow the HTA result? The current EUnetHTA guideline draft for therapeutic medical devices lines out: "The short time frame and regulatory landscape limit the performance of randomized controlled trials with sufficient sample size and follow-up. Results may already be outdated when finally available and a new model of a product may be introduced during the course of a trial.(14) In addition, the reference technology is also subject to modification.(15) The need for new clinical studies for small modifications is unclear.(16) Similarity of products and how to define it is not only an issue for successive modifications of a specific product but also for products of different manufacturers. The question of which devices can be grouped into one "class" (e.g., in terms of technical comparability) is important in health technology assessment for the choice of comparator in the evaluation of new technologies.(17)" Therefore, the guidelines in tables A1, A13 most probably refer to the assessment of pharmaceuticals or at least were written with pharmaceuticals in mind or need to be rethought to better reflect the properties and the market access reality of non-drug technologies. Conducting a systematic review takes time that decision makers and affected patients may not have. HTA in Europe should consider these stakeholder groups as customers and should		recommendation is merely a recommendation and not a requirement. If it could be justified why a systematic review would not be useful, we do not see that this guideline would force anyone to do it.

			strive to better serve their needs.		
26	9	182	Please provide references that comply with the request of reporting details within the usual restrictions of amount of content of journals.	☒ major ☐ minor ☐ linguistic	The references can be found in each section we refer to. Please, see also table A!.
27	9	182, recomm. No 4	– It is unclear if EUnetHTA means by "model" populated model or model structure. – It is unclear what EUnetHTA means by "made available." It is unclear to whom it should be made available. – In our opinion it would be acceptable if an electronic version is made available to payers/assessors in confidence. – In our opinion it would be unacceptable if an electronic version is made available to the public.	☒ major ☐ minor ☐ linguistic	We have slightly modified the recommendation to "Providing an electronic version of the model to users could enhance the transparency and usefulness further."
28	9	182, recomm. No 5	– It would be valuable to have a summary of valuation methods of indirect costs (e.g. human capital method or friction cost method) used in different EUnetHTA member states. – In our opinion the value of this document will be increased if EUnetHTA provides a recommendation on the valuation methods of indirect costs. – We agree that the societal perspective should be part of the sensitivity analysis. – It would be valuable if EUnetHTA could clarify whether it should be part of the base case analysis.	☒ major ☐ minor ☐ linguistic	Hardly any guidelines state how indirect costs should be estimated. Therefore, it is beyond the scope of this document to provide such a recommendation, even though we agree that it would be very valuable.
29	10	182, Recomm. No 7	– It would be valuable if EUnetHTA could clarify the use of the QALYs further. – In our opinion a broader perspective should encompass the utility of caregivers as well as patients (e.g. other's quality of life's impacted by an intervention). – In our opinion QALYs should be possible to have	☒ major ☐ minor ☐ linguistic	As there are different opinions about the use of QALY within the EUnetHTA members, we cannot be clearer about its use.

			differentiated weights for levels of utility to reflect equity considerations.		
30	10	182	If there is an intended difference in meaning of sentence one and two of recommendation 9, it could be formulated more clearly.	☐ major ☐ minor ☒ linguistic	Thank you for your advice! The sentences were rephrased, though not changed in content. „In a CEA or CUA, results should be presented in terms of absolute and incremental values, separately for both costs and health outcomes and in terms of incremental cost-effectiveness ratios (ICERs)."
31	11	191	Suggest replacing the term 'composition of economic evaluations" with the "features of an economic evaluation."	☐ major ☐ minor ☒ linguistic	We have revised accordingly.
32	11	199-200	Surgical intervention, medical intervention, complex intervention etc. should be added here and differentiated from or included in the term "treatment".	☒ major ☐ minor ☐ linguistic	We have revised accordingly.
33	11	199-200	– *Replace:* "various" with "more". – *Suggest removing to*: "diagnostic, preventive technology or a treatment".	☐ major ☐ minor ☒ linguistic	We have revised this sentence.
34	11	204	Perhaps it would be worth specifying here that the comparator may not actually be an active treatment. Throughout the document there appears to be an assumption that it is one (drug) treatment compared with another (drug) treatment but these guidelines would apply wider than that.	☐ major ☒ minor ☐ linguistic	We have added a line to make this clearer.

35	11	206	The title refers to "indirect comparisons" but descibes in the text that follows both indirect treatment comparisons and mixed treatment comparisons (direct and indirect evidence). The term "network meta-analyses" would be more appropriate here to cover the paragraph presented. Recommend considering the ISPOR taskforce studies for network meta-analyses to check the specification here.	☒ major ☐ minor ☐ linguistic	Thank you for this advice. However, we chose to stay with the term "indirect comparisons" as it is the use of indirect comparisons that is of interest here. It is also more commonly understood than the more technical word "network meta-analysis".
36	11	207-208	– *Replace:* "The identified studies" with "These studies".	☐ major ☐ minor ☒ linguistic	We have revised accordingly.
37	11	209	– *Replace*: "relative effectiveness" with "relative efficacy and relative effectiveness".	☒ major ☐ minor ☐ linguistic	We have revised accordingly.
38	11	210	Indirect comparisons can infer relative effectiveness but only if indirect comparisons are both technically feasible and clinically plausible – perhaps that caveat should be added rather than saying these can always be done	☐ major ☒ minor ☐ linguistic	We have strived to keep the text quite simple and made no change at this place.
39	11	213	Should „relevant" be „necessary"?	☐ major ☐ minor ☒ linguistic	Thank you for this comment. However, we have chosen not to change since it may not always be necessary.
40	11	219-220	Add „…and should be the most relevant to the decision making."	☐ major	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.

				☐ minor ☒ linguistic	
41	11	223-224	– *Replace:* "no matter on whom these costs and consequences fall" with "regardless of who is responsible for these costs and consequences".	☐ major ☐ minor ☒ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
42	11	249	Add "on the intervention that provides" between "…information on" and " the 'greatest effect…".	☐ major ☐ minor ☒ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
43	11	266	Add after benefit „in money-value"?	☐ major ☐ minor ☒ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
44	12	223	Maybe worthwhile to indicate that the length of time horizon depends on the expected length of the treatment cycle (In addition to the economic evaluation).	☐ major ☒ minor ☒ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
45	12	233-234	– Could EUnetHTA please clarify what is meant by "future generations".	☐ major ☐ minor ☒ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
46	12	242	– *Suggest to reword:* "commensurable or not" to "commensurable or not and how results are expressed."	☐ major	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.

				☐ minor ☒ linguistic	
47	12	259	– QALY should be written out as Quality Adjusted Life Years (QALYs) as it is the first time QALYs are mentioned in the document.	☐ major ☐ minor ☒ linguistic	We have revised accordingly.
48	12	265	– *Delete* „the production of the".	☐ major ☐ minor ☒ linguistic	We have revised accordingly.
49	13	307-310	– Indirect costs can also include reduced productivity while being at work (so called presenteeism). – *Add* specific definition for indirect costs. We propose to define indirect costs as "resources forgone as a result of a health condition" (reference: http://www.hsph.harvard.edu/obesity-prevention-source/obesity-consequences/economic/#references)	☒ major ☐ minor ☐ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it. Furthermore we think presenteeism is covered by this sentence "…reduced working capacity due to illness and disability".
50	14	315	– *Replace:* "The type of analysis chosen and outcomes measure" with "The selection".	☐ major ☐ minor ☒ linguistic	We have revised accordingly.
51	14	323	– Put Quality adjusted life years in the title, to be consistent with other headings (compared to line 277).	☐ major ☐ minor ☒ linguistic	We have revised accordingly.

52	14	330	Nothing has been said about how QALY is calculated with QoL so far.	☐ major ☒ minor ☒ linguistic	We have added an extra sentence to improve this. "The quality of life (QoL) aspects of the QALY are captured in a QoL weight."
53	14	330	Should this be HRQL rather than QoL?	☐ major ☐ minor ☒ linguistic	Thank you. However, we choose not to change this.
54	14	332	I would not consider the use of a VAS as a method in itself – it's a tool which can be used in a number of different ways	☐ major ☒ minor ☐ linguistic	We choose not to change this since it is usually described as a method in the literature.
55	14	334	The use of the word patients is not correct – I believe you can use those questionnaires with non-patients	☐ major ☒ minor ☐ linguistic	We have revised accordingly.
56	14	347	I do not believe that surrogate endpoints can be used as a substitute – they can be used as a proxy for final outcomes but that is not the same premise. Also, if you are to use a surrogate endpoint, you should demonstrate the associated of that endpoint with the final outcome of interest (eg how LDL levels are associated with cardiovascular events)	☒ major ☐ minor ☐ linguistic	We have changed "substitute" to "proxy".
57	15	355-356	– *Suggest to reword to*: "There are three general approaches to express benefits in monetary units (as in CBA);"	☐ major ☐ minor	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.

				☒ linguistic	
58	15	362-363	– *Suggest to reword to*: "Results of the economic evaluation should be presented in accordance with the economic evaluation used"	☐ major ☐ minor ☒ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
59	15	379	Have not defined what is cost-effective yet.	☐ major ☒ minor ☐ linguistic	The definition of cost-effectiveness has been moved to before "Results of the economic evaluations".
60	15	384	EVPI should be a subset of EVI do not needed in header. Not sure this section has enough information	☐ major ☒ minor ☐ linguistic	We have revised the heading. This section is very short but that is intendedly as EVPI (EVI) is not often covered in the various guidelines. However, we have added an extra line about EVI.
61	15	384-388	Why not describe EVI then EVPI as a special case? Also the concept of value of information and uncertainty has not been introduced, so this part does not come naturally here.	☐ major ☒ minor ☐ linguistic	This text has been moved to the section about uncertainty. See also reply on comment above.
62	15	389	The title should be NMB and NHB in terms of the contents.	☐ major ☐ minor ☒ linguistic	We have revised this accordingly.
63	15	389	Title should include both NMB and NHB	☐ major ☒ minor	We have revised this accordingly.

				☐ linguistic	
64	15	390 & 391	The section "net health and the net monetary benefit (NHB and NMB) are a framework to display uncertainty in cost-effectiveness analysis (15, 16) and it can also" could be replaced by "net health benefit and the net monetary benefit -approaches (NHB and NMB) provide a framework to display uncertainty in cost-effectiveness analysis (15, 16) and they can also"	☐ major ☐ minor ☒ linguistic	We have revised this accordingly.
65	16	396	We suggest to clarify the title as "Cost-effectiveness threshold"	☐ major ☐ minor ☒ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
66	16	396	The section on the cost-effectiveness threshold should distinguish between the theoretical notion of a cost-effectiveness threshold and the practical implementation of cost-effectiveness thresholds in different countries. This is an area of great controversy in the literature and practice where many differences exist between countries with respect to system level objectives, decision processes and how other factors are taken into account alongside cost-effectiveness evidence. Suggest that this is made clear under the cost-effectiveness threshold concept.	☒ major ☐ minor ☐ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
67	16	410-411	Making decision criteria more explicit and aligned across Europe could attract more investments in healthcare because of waste reduction on the market acces pathway.	☒ major ☐ minor	We agree.

				☐ linguistic	
68	16	413-431	Could this paragraph could be organized better to reflect 1) modeling clinical decision making process (decision tree etc), 2) modeling outcomes given decisions (state transit and discrete event models etc), and 3) evaluating outcomes (direct calculation or simulation)	☐ major ☒ minor ☐ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
69	16	413-431	– Markov decision processes are lacking.	☐ major ☒ minor ☐ linguistic	This text is taken from the ECO domain of the HTA Core Model, so we choose not to change it.
70	16	435	In "Discounting, i.e., calculating …", should "i.e.' be "in"?	☐ major ☐ minor ☒ linguistic	We have deleted ",".
71	16	427	ISPOR has published a series of guidelines (http://www.ispor.org/GuidelinesIndex/Default.aspx#HEEM)	☒ major ☐ minor ☐ linguistic	We are aware of these. The most recent and relevant guidelines has been added to the list of related documents. A link to the ISPOR guidelines index has also been provided.
72	17	471	– *Suggest to reword to*: "Multi-way sensitivity analysis".	☐ major ☐ minor ☒ linguistic	We have revised this accordingly.
73	18	478	Threshold analyses can also be referred to as tipping point analyses in the statistics literature	☐ major ☒ minor	We have decided to keep the term threshold analysis since the term tipping point is less frequently used in this context.

				☐ linguistic	
74	18	482	I'm not sure the notion of non-linear model is adequate in this context. The definition of Markov models and Monte-Carlo simulations should be revisited.	☒ major ☐ minor ☐ linguistic	We agree with you and have revised this section briefly.
75	18	483	The concept of non-linear models has not been introduced before this point so needs some discussion and context	☐ major ☒ minor ☐ linguistic	See previous comment
76	18	483-484	Is this statement correct? PSA is required to evaluate the impact of uncertainties when it is considered random, which is often considered more appropriate than fixing a parameter at one value in DSA.	☒ major ☐ minor ☐ linguistic	See previous comment
77	18	490	– It is unnecessary to mention the number of 1,000 as there is no scientific rationale for this.	☐ major ☒ minor ☐ linguistic	The number is just used as an example to show that it should be many times (not just a few). Another example has been added.
78	19	494	Suggest using "Decision problem" rather than "Problem statement" because this is a more widely used term and can be found in the main health economics text such as Drummond et al. 1997. Drummond, M., Sculpher, M., Torrance, G., O'Brien, B. & Stoddart, G. (2005). Methods for the Economic Evaluation of Health Care Programmes. 3rd ed. New York: Oxford Press. 289-	☐ major ☐ minor ☒ linguistic	This heading is predefined in this type of EUnetHTA document and is the title for the problem statement of the guideline, not a health economic analysis.

			290.		
79	19	514	"Nevertheless, if economic evaluations performed within EUnetHTA are to be useful outside of the authors' own country, it is essential that the methodology reflects the general view of the EUnetHTA members or that the effect of using different methods is explored in sensitivity analyses." This sentence reflects the three key failings: 1) That there is no single methodology which reflects all EUnetHTA countries 2) That the recommendation – to then do all other types of analyses as sensitivity analyses – will make submissions unduly cumbersome and therefore not add value. In the worst case scenario, a submission would have to include a full analysis for each additional country as a sensitivity analysis to be handled by the submission country – which would not be skilled in evaluating using methodology not natively adopted. 3) There is only a reflection on methodology, not on baseline settings. Any analysis which does not take the national clinical and economic setup into consideration is bound to be irrelevant.	☒ major ☐ minor ☐ linguistic	This document is intended as a guideline when EUnetHTA-members make their own CEA, not as a guideline for submissions to specific reimbursement agencies. HTA is not only conducted to assess pharmaceuticals and devices but also on many procedures such as surgery, rehabilitation etc. We would argue that if an analysis is transparent enough, it is often possible to adjust for national clinical and economic setups. However, we have decided to remove these lines as they may cause some confusion.
80	20	577	REA of pharmaceuticals	☒ major ☐ minor ☐ linguistic	We have revised accordingly.
81	21	580	Clincial evidence assessments of non-drug technologies need to be specified in the context of the above mentioned medical device guideline before it can be agreed to that this health	☒ major ☐ minor	Thank you. However, this comment is beyond the scope of this report.

			economics guideline applies to non-drug technologies.	☐ linguistic	
82	21	584	The ISPOR modelling taskforce paper could also be stated here: "Pitman et al. Dynamic Transmission Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-5. Value in health 15 (2012) 828–834"	☒ major ☐ minor ☐ linguistic	Thank you, this reference has been added.
83	21	584	it may be worthwhile pointing out there are many other EMA CHMP guidance documents other than the draft guideline on subgroup analyses that could be referenced and may be relevant for the EUnetHTA guideline, there is a list here: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000602.jsp&mid=WC0b01ac05807d91a4	☐ major ☒ minor ☐ linguistic	The most recent and relevant documents are now mentioned and we have also included a link to EMA's website.
84	21	585-598	ISPOR: Good Research Practices for Cost-Effectiveness Analysis Alongside Clinical Trials: The ISPOR RCT-CEA Task Force Report: Ramsey S1, Willke R, Briggs A, Brown R, Buxton M, Chawla A, Cook J, Glick H, Liljas B, Petitti D, Reed S.: Value Health. 2005 Sep-Oct;8(5):521-33. Applying Dynamic Simulation Modeling Methods in Health Care Delivery Research – The SIMULATE Checklist: An ISPOR Simulation Modeling Emerging Good Practices Task Force Report: http://www.ispor.org/TaskForces/Simulation-ModelingApps-HCDelivery.asp Principles of Good Practice for Budget Impact Analysis: Report of the ISPOR Task Force on Good Research Practices – Budget	☒ major ☐ minor ☐ linguistic	The most recent and relevant documents are now mentioned and we have also included a link to ISPORs website.

			Impact Analysis: Josephine A. Mauskopf, PhD,1 Sean D. Sullivan, PhD,2 Lieven Annemans, PhD, MSc,3 Jaime Caro, MD,4 C. Daniel Mullins, PhD,5 Mark Nuijten, PhD, MBA, MD,6 Ewa Orlewska, MD, PhD,7 John Watkins, RPh, MPH,8, Paul Trueman, MA, BA9 Principles of Good Practice for Budget Impact Analysis: Report of the ISPOR Task Force on Good Research Practices—Budget Impact Analysis http://www.ispor.org/budget-impact-health-study-guideline.asp: Volume 10 • Number 5 • 2007 VALUE IN HEALTH http://www.ispor.org/taskForces/TFindex.asp Health Economic Evaluation Publication Guidelines – CHEERS •Measuring Drug Costs in CEA: Issues and Recommendations •Measuring Drug Costs in CEA: A Societal Perspective •Measuring Drug Costs in CEA: A Managed Care Perspective •Measuring Drug Costs in CEA: Medicare/Medicaid Perspective •Measuring Drug Costs in CEA: An Industry Perspective		

			•Measuring Drug Costs in CEA: An International Perspective •Medical Nutrition Products - Outcomes Research (in development) •Quality Improvement of Cost Effectiveness Research •Transferability of Economic Evaluations Across Jurisdictions •Conceptualizing a Model: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-2 •Dynamic Transmission Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-5 •Modeling Good Research Practices - Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1 •Modeling Studies •Modeling using Discrete Event Simulation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-4 •Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6 •Model Transparency and Validation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-7		

			•Simulation Modeling Applications in Health Care Delivery Research - Emerging Good Practices Task Force •State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-3 •Conjoint Analysis Applications in Health Good Research Practices •Conjoint Analysis Experimental Design Good Research Practices • Conjoint Analysis - Statistical Analyses (in development) •Health State Utility Values – Mapping for Cost per QALY Economic Analysis (in development) •Health State Utility Values – Measurement for Economic Models in Clinical Studies (in development) •Moving the QALY forward - Consensus development Modeling Study Questionnaire for Health Care Decision Making Use of Outcomes Research in Decision-making Quantitative Risk-Benefit Methods for Assessing Drug Safety and Efficacy: Report of the ISPOR Risk-Benefit Management Working Group: Guo JJ, Pandey S, Doyle J, et al. A review of quantitative risk-benefit methodologies for assessing drug safety and efficacy – Report of the ISPOR Risk-Benefit Management Working Group.		

			Value in Health 2010; 13(5):657-666.		
85	21	605-606	Pharmaco-economic guidelines do not necessarily apply to other technologies.	☒ major ☐ minor ☐ linguistic	Other sources were used as well. ISPOR's database was only a starting point.
86	21	621	Pharmaco-economic guidelines do not necessarily apply for other technologies.	☒ major ☐ minor ☐ linguistic	Other sources were used as well. ISPOR's database was only a starting point.
87	23	666	Not clear what is meant by the sentence 'NICE in England has three different guidelines for different types of technologies', from the annex tables I assume they mean medical devices vs pharmaceuticals vs public health. Not particularly clear. You then also need to be clear throughout, which process you are referring to because it can get very confusing just saying England has 'x' if it is not in the 'standard' NICE guidelines/reference case.	☐ major ☐ minor ☒ linguistic	We refer to the guidelines for the technology appraisals that cover all type of technologies, specific guidelines for diagnostics and specific guidelines for medical devices (See table A1). This has been clarified.
88	24	674	No vaccine evaluations have been looked into => cf to comment n°1	☒ major ☐ minor ☐ linguistic	We fully agree that there are specific issues in guidelines on vaccines, eg, dynamic models etc. However, there are very few guidelines for health economic evaluations solely pertaining to vaccines and none of the contact persons of the EUnetHTA members have provided information about specific guidelines for vaccines, even though we have asked for guidelines for all different types of health technologies. It is also our understanding that the JCVI code of practice cannot be understood as a guideline concerning

					health economic evaluations and that it refers to other NICE guidelines on these issues. However, we have, based on comments from the reviewers, decided to, as an example, include additional information from one of the guidelines (France) concerning discounting in relation to the evaluation of public health programmes such as vaccines.
89	24	675-679	Please add the full references including download link or a zip file with all included guidelines: Does a positive result applying these methods also lead to a positive decision?	☒ major ☐ minor ☐ linguistic	All guidelines are listed in table A1 in Annexe 3. This has now been clarified in the text. This document is intended as a guideline when EUnetHTA-members make their own CEA, not as a guideline for submissions to specific reimbursement agencies.
90	24	690	The documents states „routine" clinical practice but many of the issues faced are around how you define „routine" (most frequently used, determined by guidelines etc)– perhaps that needs more discussion	☐ major ☒ minor ☐ linguistic	We believe that this is clarified in the examples and the summary.
91	24	700-702	The impact of (country-specific) value judgements within the assessment and the appraisal phase should be considered to be investigated.	☒ major ☐ minor ☐ linguistic	This could perhaps be a subject for future collaborations.
92	24	707	Please define "extendedly dominated" in comparison to "dominated".	☐ major ☒ minor ☐ linguistic	This is a common term, for example presented by Drummond et al. As it requires extensive explanations, we choose not to go deeper into this in the document.

93	25	725-728	The choice of comparator (type of hypothesis as well) also depends on the goal of the assessment, e.g. price premium or price parity.	☒ major ☐ minor ☐ linguistic	This is true, but this information is not available in the regional guidelines.
94	25	736	Could we emphasize the point that RCTs cannot be done in all possible comparators used in standard of care and this is a challenge when trying to select a comparator for a confirmatory trial, which comparator in an international trial run in multiple countries best reflects standard of care	☒ major ☐ minor ☐ linguistic	The text in the document (as well as the conclusions) are based on the extractions of all guidelines.
95	25	737-741	There are likely to be difficulties in undertaking analyses which would attempt to be representative of all the countries involved in EUnetHTA in terms of a relevant comparator/ relevant comparators. Therefore, I would suggest the following modified wording, e.g.,: "Based on the results of the current review of the guidelines used by EUnetHTA members and previous EUnetHTA guidelines, it is recommended that the comparator(s) should reflect the most relevant alternative treatment(s) used in clinical practiceOther relevant comparators should also be considered. The choice of comparator(s) should be clearly presented and clearly justified." or "Based on the results of the current review of the guidelines used by EUnetHTA members and previous EUnetHTA guidelines, it is recommended that the comparator(s) should reflect the	☒ major ☐ minor ☐ linguistic e	We have revised this according to the first suggestion.

			most relevant alternative treatment(s) used in clinical practice, especially any comparators relevant in the countries involved in EUnetHTA. Other relevant comparators should also be considered. The choice of comparator(s) should be clearly presented and clearly justified."		
96	25	739	Statement is about „most relevant alternative treatments" but this does not allow for the fact that those may be „watch and wait" surveillance, best supportive care etc. I feel this is too narrow as a definition.	☐ major ☒ minor ☐ linguistic	Treatment has been replaced with intervention. However, If "watch and wait" etc is the more relevant alternative, it is certainly included in this statement. This could also be "no treatment" if this is the most relevant alternative.
97	25	738-740	– It would be valuable to know what sources in general (or examples thereof) could be considered as appropriate to identify such comparators if treatment guidelines are unavailable – Can EUnetHTA please comment on how the standard of care should be derived and evaluated if it is different across EUnetHTA member states?	☒ major ☐ minor ☐ linguistic	The text is based on the extractions of information from the guidelines.
98	25	752-754	It needs to be made clear for which types of technologies this type of assessment is applied and how often to delinate how important this guidance really is used for decision making.	☒ major ☐ minor ☐ linguistic	We have rephrased this sentence slightly but it is beyond the scope of this project to write about how often HE-analyses based on the different guidelines were performed. The text applies to all kind of technologies though it was written with (new) drugs in mind.
99	26	767	Any words on multiplicity and how to conduct such analyses? Should we add a sentence about statistical and clinical rational ("Justification") for the subgroups? I would favour more stringent guidelines on the best practices when running such subgroup analyses in the context of health economics	☐ major ☒ minor ☐ linguistic	We agree with your comment, however, most of the guidelines were not more specific about this issue.

			evaluations		
100	26	768	Subgroups are often requested by national authorities with no direct agreement between authorities. Therefore, this suggestion is unhelpful and disregards current practice.	☒ major ☐ minor ☐ linguistic	However, this conclusion is based on the extraction of data from the guidelines.
101	26	770	– *Replace:* "is believed" with "has a clinical rationale"	☐ major ☐ minor ☒ linguistic	We have revised accordingly.
102	26	771	The choice of subgroups needs to be clearly justified and described I agree but they also need to be clinically plausible and that point does not come across in this sentence	☐ major ☒ minor ☐ linguistic	We agree and have, based on this and other comments in the review, tried to clarify this.
103	26	769-771	– In our opinion it should be stated that the biological/medical rationale should be used to justify the choice of subgroups, not purely based on cost-effectiveness. – In our opinion it should be added that the post-hoc choice of subgroups is problematic due to the higher rate of false positive findings (i.e. identify a spurious difference between groups). The scientific basis for post-hoc sub-grouping and benefit this brings to patients is questionable. – Reference: Ruof et al. (2014), Questioning Patient Subgroups for Benefit Assessment: Challenging the German Gemeinsamer Bundesausschuss Approach, in: Value in Health 17 (2014), p. 307-309. (http://dx.doi.org/10.1016/j.jval.2014.05.001)	☒ major ☐ minor ☐ linguistic	We agree on the point concerning biological/ medical rationale and have tried to clarify this in the recommendation. See also comment 101-102.
104	26	785-	Before EUnetHTA requests reproducibility, EUnetHTA and its' partner organisations should at first publish the protocols of all	☒ major	We agree with your comment and have forwarded this to the lead partners.

		789	their reports on their website to comply to this: "the methods of the review should be reported in sufficient detail to enable the review to be reproduced." Often-even if published - HTA protocols do not comply with the criterion of sufficient detail, HTA assessors may still have discretionary choices to make during the conduct of the assessment.	☐ minor ☐ linguistic	
105	26	792-794	"Only a few of the countries with guidelines request that a systematic review over previous economic evaluations is presented (Croatia (58), England (12), France (53), Poland (59), 778 Slovakia (60) and Spain (AETSA (61)) (Table A6 in Annexe 5)." „Based on the results of the current review of the guidelines used by EUnetHTA members, it is recommended that the clinical evidence is collected by a systematic review of the literature." Why does this EUnetHTA guideline draft come to the contrary result as the majority of the guidelines? What is the justification? Can it at all be possible to draw such a contrarian conclusion? It is not surprising that the producers of systematic reviews recommend their use (this is only a statement derived from a self-assessment), but is there a real business case for HTA and for which technologies? To what extent is HTA really used for the different types of technologies across Europe and to what extent do decisions follow the HTA result? Decision makers or patients need may need quick decision based	☒ major ☐ minor ☐ linguistic	It seems like it has been a misunderstanding concerning two aspects here. Most organisations agree that the clinical effects should be collected by a systematic review and that is what the recommendation address. However, only a few say that a systematic review of former economic evaluations is necessary and therefore the conclusion about this states that it is "regarded as useful to conduct a systematic review…".

			on a timely assessment of the available evidence. How can these types of needs also be served? Conducting a systematic review especially on health economics comes along with the risk of to narrowly defining criteria and thus missing relevant publications and often is a too high burden to provide data as quick as possible to decision makers. A documented and reproducible literature search may often be sufficient. This recommendation reflects the values of the contributing EUnetHTA partners, it is necessary to clarify if the "end customer" - e.g. the decision maker- follows the same or similar values. For example the decision maker could rather prefer a quick decision on the available evidence, rather than waiting for a European full core model being conducted and locally adapted before it is presented to him.		
106	26	792-794	– *Suggest to add: "*Systematic reviews for existing models serve two purposes. (1) to consider whether the current structure and assumptions used within the economic evaluation are appropriate and/or have been used previously to help contrast against earlier methods, and (2) to compared modelled results across interventions. – In our opinion it should be stated that systematic reviews of quality of life utility studies and mapping studies are useful in determining which are the most appropriate for the evaluated population.	☒ major ☐ minor ☐ linguistic	We agree with the importance of your comment, but it is not based on the guidelines used for this document.
107	26	792-794	Could the conclusion be more specific than saying a systematic review is regarded useful?	☐ major ☒ minor	Based on that only a few guidelines say that this is necessary and many do not mention it at all, we feel it is hard to be more specific.

				☐ linguistic	
108	26	794	If this is a new technology which is not yet on the market, a lit search of economic evaluations for the technology will identify close to 0 published studies. Would it not be better to do a search for economic evaluations of the disease area and / or drug class (where that is relevant)? It should also be "Recommended" instead of "regarded as useful".	☒ major ☐ minor ☐ linguistic	Based on that only a few guidelines say that this is necessary and many do not mention it at all, we feel it is hard to be more specific.
109	27	803	"this may mean a life-time" may be better stated as something like "this may mean estimating costs and outcomes for the estimated remaining life time of the patients"	☐ major ☐ minor ☒ linguistic	This has now been revised.
110	27	803-805	– In our opinion it should be stated more clearly that the German interpretation of economic evaluation and subsequent time horizon is considerably different than many of the guidelines referenced from the other countries.	☒ major ☐ minor ☐ linguistic	However, we think this statement „…The only guideline which partly depart from this view is one from Germany" is clear.
111	27	810-811	– Please clarify what EUnetHTA means by unintended future costs.	☒ major ☐ minor ☐ linguistic	We have deleted "both intended and unintended" since this part was considered unnecessary in this context. We have also forwarded this comment to the authors of the ECO domain of the Core Model.
112	27	815-819	– *Replace:* "all important" with "all relevant" – In our opinion there is a lack of clarity in the definition of "sufficiently long". (time horizon that is typically chosen is a patient's lifetime, although shorter periods may be useful depending on the aims of the study or the chosen health	☐ major ☐ minor ☒ linguistic	We have changed this to "all relevant". "Sufficiently long" is the most commonly used word in the local guidelines.

			outcome).		
113	27	816-819	Given the dearth of information concerning the long-term costs and effects of some interventions, I suggest that the following sentence would be added to the recommendation: "It is important that the choice concerning any alternative time horizon are clearly justified and described."	☒ major ☐ minor ☐ linguistic	This line has been added.
114	27	819	Why does this section not also include a statement about exploring the impact of the choice of time horizon on the outcomes? Sensitivity analyses on the choice of the time horizon could also be performed Justification should be clearly provided	☐ major ☒ minor ☐ linguistic	This has only been mentioned in a few guidelines. As an example, we have added the following text in section 2.3.2. "Nevertheless, there are guidelines that ask for other time horizons in sensitivity analyses. For example, the Scottish guidelines (61) further specify that results (in cost per QALY gained) need to be reported at different time horizon intervals e.g. at end of study follow-up, at 5 years follow-up and at five-year intervals thereafter."
115	27	827	I think this sentence is misleading. I believe there are two different issues here – the natural history of the disease and the characteristics of the technology should indeed drive the choice of model. The availability of the data to support this is a different issue – it may be that you need to go with the „correct" model structure and then work out how best to use the data you have but I'm not sure I agree that the availability of data is a driving factor in the same way as disease and treatment characteristics.	☐ major ☒ minor ☐ linguistic	We agree with your comment, but this is what is stated in the guidelines covered by the review, not the recommendation of the EUnetHTA guideline.
116	27	836	Should „clinical effects" more clearly include side effects? I feel	☐ major	This is almost a quote from the Polish guideline, so we don't want to add more aspects to it.

<table>
<tr><td></td><td></td><td></td><td>that this concept is missing here</td><td>☒ minor
☐ linguistic</td><td></td></tr>
<tr><td>117</td><td>27 & 36 & 71</td><td>832 & 1220 & 2160</td><td>It states that guidelines in England require CEA AND CUA, this does not seem to be the case according to the NICE guidance pages; http://www.nice.org.uk/article/pmg6/chapter/7-assessing-cost-effectiveness
"
<table><tr><td>Type of economic evaluation</td><td>Cost–utility analysis with fully incremental analysis</td></tr></table>"
You also then contradict yourself in the A8 table (pg 71) – where under England it is stated:
"
Engl and CUA (Technology Appraisals and NICE Diagnostics Assessment Programme)
CCA (NICE Medical Technologies Evaluation Programme Methods Guide)
"
Implying CEA is NOT necessary.</td><td>☒ major
☐ minor
☐ linguistic</td><td>Cost per life year gained is indeed not required (although NICE requires to present LY gained). Therefore we have omitted England from the list in section 2.3.3 but not in 2.3.7.3.</td></tr>
<tr><td>118</td><td>28</td><td>848</td><td>This sentence is misleading since a QALY is not a measure of</td><td>☐ major</td><td>We have rephrased this sentence to: "…when no other measure capturing/describing quality of life</td></tr>
</table>

			quality of life so the use of the word „other" does not make sense	☒ minor ☐ linguistic	is available."
119	28	861	The use of the word „formulations" is incorrect	☐ major ☐ minor ☒ linguistic	This has been revised.
120	28	870-871	Need to explain why QALYs may not be appropriate for severe pain over short period of time – this is not straightforward.	☐ major ☒ minor ☐ linguistic	This example is from the Swedish guideline, which is not more specific than that. However, for example, if you have one day with severe pain, that would not give much impact on a QALY, but still many people would be willing to pay much money to get rid of it (for example tooth ache).
121	28, 30, 32	888-889, 973, 1023 etc.	Is the EUnetHTA Core Model 'another guideline'? I thought this guideline and the EUnetHTA Core Model were to be seen as complementary.	☐ major ☐ minor ☒ linguistic	You are right that is not a guideline but the instructions for the model contain some recommendations that need to be mentioned since the guideline and the CORE model instructions should be complementary.
122	29	892	The phrase „suitable data" is not explicit enough and may be used a justification for not collecting data which may be suitable	☐ major ☒ minor ☐ linguistic	This text comes from the Costs and Economic Evaluations (ECO) domain of the HTA Core Model®, and we would therefore prefer to not change it.
123	29	894	Include a statement about use of CBA as well in the conclusion	☐ major ☒ minor ☒ linguistic	It is hard to include one statement about CBA as there are different opinions about its use. The recommendations are only based on commonalities or transferability improvements.

124	29	894-900	– Should be put in orange box	☐ major ☐ minor ☒ linguistic	We have revised the manuscript accordingly.
125	29	895-900	– It will be more relevant to use CUA in case where the morbidity/disability has an impact on the quality of life. If so, CUA should be the first analysis and the CEA should also be included. If the disease/disability doesn't have any consequence on QoL, CEA should be preferred. – It would be valuable to know EUnetHTA's recommendation concerning the use of CBA.	☒ major ☐ minor ☐ linguistic	Our conclusions are based on what the different guidelines states. Furthermore, it is hard to include one statement about CBA as there are different opinions about its use.
126	30	932	Choice of models is highly dependent on the intervention analyzed. This is especially important for vaccines evaluations for which transmission needs to be modelled.	☐ major ☒ minor ☐ linguistic	We agree with your comment, but this text only refers to what the different guidelines recommend.
127	30	935	Budget-impact-models	☐ major ☒ minor ☐ linguistic	Budget-impact-models are rarely mentioned by the guidelines. Often, it is rather defined as another type of analysis.
128	30	955	From our understanding, French guidelines ask for three analyses with three different product's prices but this does not correspond exactly to optimistic, pessimistic and neutral scenarios	☐ major ☒ minor ☐ linguistic	In this paragraph which described scenarios of extrapolation, the French guidelines ask for 3 scenarios when extrapolating data. It is true that when submitting dossiers, manufacturers are requested to provide scenarios with 3 different prices but this is a different issue that it is not described in the HAS guidelines for economic evaluation.

129	31	978	Should the model also be calibrated	☐ major ☒ minor ☒ linguistic	This text comes from the Costs and Economic Evaluations (ECO) domain of the HTA Core Model®, and therefore we don't want to change it.
130	31	997	CatSalut (Catalonia-Spain) considers the healthcare perspective in the main analysis. Societal perspective can be developed in a complementary analysis.	☒ major ☐ minor ☐ linguistic	Spain was removed from the parenthesis as several perspectives are recommended.
131	31	1003 - 1007	Why does a collective perspective including all costs mean to exclude indirect costs?	☒ major ☐ minor ☐ linguistic	Economic evaluation at HAS is based on the analysis of production costs. Consequently, only direct costs are taken into account in reference case analysis, and included in the incremental cost-effectiveness ratio. An analysis of the indirect costs, if considered relevant by the author of the study, is presented in an additional analysis.
132	31	1008	Outcomes considered for vaccines evaluations may also be wider as they should include benefits for a wider population which may be protected through indirect effect even if not directly vaccinated (herd immunity)	☐ major ☒ minor ☐ linguistic	We fully agree that there are specific issues in guidelines on vaccines, eg dynamic models etc. However, there are very few health economic guidelines specifically and solely pertaining to vaccines and none of the contact persons of the EUnetHTA members have provided information about specific guidelines for vaccines even though we have asked for guidelines for all different types of health technologies. Therefore, we will not include them, but mention them. However, we have decided to include the JCVI code of practice from UK and comment on this guideline when it differs from the other UK documents.

133	31	1015	Please explain the difference between patients and individuals, e.g. family members.	☐ major ☒ minor ☐ linguistic	We have tried to clarify this. Individuals may be both patients and non-patients. We have tried to stay with the terms used in the guidelines we refer to.
134	31	1015	The phrase „recommended to include" needs modified (here and throughout the document)	☐ major ☐ minor ☒ linguistic	This has been revised.
135	32	1030	This conclusion does not actually present a conclusion per se. It rather lists an intention – and with all societal perspectives being dependent on the relevant societies, the analysis would have to be re-run for each country. The relevance across countries is not apparent and the workload likely to increase for industry and authorities alike.	☒ major ☐ minor ☐ linguistic	Due to the different views in the guidelines, it is hard to be more precise than this.
136	32	1038	This section should really be split into two different sections – on e dealing with estimating resource use and one dealing with estimating unit costs. The issues are different around data identification and transferability and they get mixed up in the current section. Many guidelines (e.g. NICE) ask for these to be done separately anyway	☒ major ☐ minor ☐ linguistic	This and the following section have been partly rewritten to make them clearer.
137	32	1060	The line is correct but the actual 'standard' NICE methods guidance states; "Value added tax (VAT) should be excluded from all economic evaluations, but included in calculation of the budgetary impact when the resources in question are liable for this tax."	☐ major ☒ minor ☐ linguistic	This information has now been added.

138	33	1068	As in note 8: there is no true alignment on societal focus and/or costs	☒ major ☐ minor ☐ linguistic	Due to the different views in the guidelines, it is hard to be more precise than this.
139	33	1069 - 1071	– It should be stated that the base case analysis should include direct medical costs but mandatory sensitivity analyses should include direct non-medical costs and then the inclusion of indirect/societal costs.	☒ major ☐ minor ☐ linguistic	We have rephrased the sentence to: "… there is an apparently plain consensus that all direct costs should be included in the main analysis. It is also recommended to present indirect costs – when it is relevant - in an additional analysis"…. However, we have not written that the sensitivity analyses are mandatory.
140	33	1085	This is an example where the Scottish guidelines talk about estimating resource use but the sentence is in a section about costs – it could be read that Scotland are happy with clinicians estimating the unit cost of a resource which is not the case	☒ major ☐ minor ☐ linguistic	Related to comment 136. This section has been partly rewritten to make it clearer.
141	33	1089	The idea is interesting but does not reflect the true ability to compare as natural units can become a complicated measure in chronic diseases – and it still does not reflect systemic or clinical differences among countries.	☒ major ☐ minor ☐ linguistic	We agree with you but still think this recommendation is of value, even though there still would be several problems to transfer the analyses to other countries/settings.
142	34	1125 - 1130	The conclusion that a systematic review should be the basis is the clinical effectiveness review does not comply with the current practice as displayed in table A13: 11 of 21 guidelines accept RCTs as the basis without a systematic review. It is not surprising that the producers of systematic reviews recommend their use (this is only a statement derived from a self-assessment), but is there a real business case for HTA and for which technologies? To what extent is HTA really used for the	☒ major ☐ minor ☐ linguistic	Our recommendation is based on this line „the majority of the countries with guidelines state that they prefer systematic reviews and meta-analyses". Furthermore, none is against it. Even though many do accept RCTs without a systematic review, this does not mean that they recommend to not do the review. We have rephrased the conclusion a bit to make it clearer.

			different types of technologies across Europe and to what extent do decisions follow the HTA result? The current EUnetHTA guideline draft for therapeutic medical devices lines out: "The short time frame and regulatory landscape limit the performance of randomized controlled trials with sufficient sample size and follow-up. Results may already be outdated when finally available and a new model of a product may be introduced during the course of a trial.(14) In addition, the reference technology is also subject to modification.(15) The need for new clinical studies for small modifications is unclear.(16) Similarity of products and how to define it is not only an issue for successive modifications of a specific product but also for products of different manufacturers. The question of which devices can be grouped into one "class" (e.g., in terms of technical comparability) is important in health technology assessment for the choice of comparator in the evaluation of new technologies.(17)" Therefore, the guidelines in tables A1, A13 most probably refer to the assessment of pharmaceuticals or at least were written with pharmaceuticals in mind or need to be rethought to better reflect the properties and the market access reality of non-drug technologies. Conducting a systematic review takes time that decision makers and affected patients may not have. HTA in Europe should consider these stakeholder groups as customers and should strive to better serve their needs.		

			Conducting a systematic review comes along with the risk of to narrowly defining criteria and thus missing relevant publications and often is a too high burden to provide data as quick as possible to decision makers. A documented and reproducible literature search is sufficient. This recommendation reflects the values of the contributing EUnetHTA partners, it is necessary to clarify if the "end customer"= the decision maker follows the same or similar values. For example the decision maker could rather prefer a quick decision on the available evidence, rather than waiting for a European full core model being conducted and locally adapted before it is presented to him.		
143	34	1123	– *Replace: "*clinical effectiveness" with "clinical efficacy/effectiveness" as at the time of preparation it is unlikely to be any effectiveness data available.	☒ major ☐ minor ☐ linguistic	We choose to keep the term "clinical effectiveness" even though we agree that most of the times only its efficacy has been studied. Yet, effectiveness can be understood as the overall concept.
144	34	1128	This sentence should read „which in most cases is considered to be"	☐ major ☐ minor ☒ linguistic	We have revised this accordingly. We have added a few words to clarify.
145	34	1128	Lower-level evidence is not clearly defined – Would it be "Lower-quality" instead?	☐ major ☒ minor ☐ linguistic	We have revised to "other sources".
146	34	1128	– Real World/registry Data should be considered as an equally relevant source of clinical efficacy/effectiveness (depending	☒ major	We agree with your comment in theory, but the conclusions must be based on all the guidelines

		- 1130	on the disease area). – We suggest removing text indicating that RCT data is higher in the hierarchy than Real World Data. This is not always the case and HTAs want to make decisions based upon what occurs in real life (not a controlled setting). – Reference: Rawlins, M. (2008), De Testimonio: on the evidence for decisions about the use of therapeutic interventions, in: Clinical Medicine 8 (6), pp. 579-588. (doi: 10.7861/clinmedicine.8-6-579)	☐ minor ☐ linguistic	being used within EUnetHTA. However, a few modifications have been made to tone down the focus on RCTs.
147	34	1129 - 1130	– In our opinion the guidance would be made more useful if EUnetHTA could provide suggestions as to which criteria or checklists are recommended to assess and report the quality of all sources. – It would be valuable to know EUnetHTA's recommendation on methods for 'grading' publications on RCT.	☒ major ☐ minor ☐ linguistic	This is unfortunately beyond the purpose of this document. However, we agree that it would be useful if EUnetHTA would issue this kind of recommendation.
148	34	1139	CatSalut also accepts meta-analysis and indirect comparisons when there are not direct comparative studies.	☐ major ☒ minor ☐ linguistic	We have revised this accordingly.
149	36	1207	QALYs are not a disease specific aggregate outcome measure so the use of the word „other" is incorrect	☐ major ☒ minor ☐ linguistic	We have revised this accordingly.
150	36	1210	Inconsistent use of language – TTO and SG are not instruments but are described as such	☐ major ☒ minor ☐ linguistic	We have revised to "…certain methods such as TTO or SG".
151	36	1208	Please give the reason why the QALY should not be used across	☒ major	This is according to the German guideline referred to and not a recommendation in the

		- 1211	disease areas. Please detail the ethical and methodological concerns mentioned.	☐ minor ☐ linguistic	guideline. The German position has been clarified.
152	36	1218	This should be reworded – the main objective is to have an effect on life expectancy so the focus is not on quality of life but this does not mean that the intervention does not have an impact on HRQL (it's just not the focus)	☐ major ☒ minor ☐ linguistic	We have revised this accordingly.
153	36	1218	This should be reworded – the main objective is to have an effect on life expectancy so the focus is not on quality of life but this does not mean that the intervention does not have an impact on HRQL (it's just not the focus)	☐ major ☒ minor ☐ linguistic	Duplicate, see 152.
154	38	1285 - 1286	Please list the criteria for validated surrogates stipulated in the IQWiG's General Methods 4.1 here.	☒ major ☐ minor ☐ linguistic	Even though we agree that this could be interesting we believe that the following lines are enough in this context: „The current methodological literature frequently discusses correlation-based procedures for surrogate validation, with estimation of correlation measures at both study and individual level. IQWiG's guideline (56) on benefit assessments do therefore give preference to validations on the basis of such procedures." Interested readers are recommended to read the original document.
155	38	1305	Important note, but there is disagreement in some disease areas – such as diabetes – as to what surrogate endpoints are relevant. This complicates comparisons	☒ major ☐ minor ☐ linguistic	This is a very important point, but we do not think the conclusion needs to be revised since it states that the relationship between the intermediate/surrogate outcomes and final outcomes should be demonstrated.
156	38	1307	– In our opinion the guidance would be made more useful if EUnetHTA could provide suggestions or standards as to what	☒ major	We agree. However, very few guidelines even mention this.

			the criteria for proving this relationship should be.	☐ minor ☐ linguistic	
157	38	1307 - 1308	Expectations and methods regarding "relationship to final outcome measures, in terms of morbidity and mortality, is demonstrated." should be lined out here. List of surrogate outcomes accepted by involved HTA partners should be attached as an appendix.	☒ major ☐ minor ☐ linguistic	This is beyond the scope of this document since very few guidelines even mention this.
158	38	1310	WTP measures the value of outcomes. It seems the distinction between an outcome measure and the value of outcomes not very clear in some places.	☒ major ☐ minor ☐ linguistic	We have tried to make this clearer.
159	38-39	1311 - 1342	WTP is not necessarily an outcome measure, it is needed to define a threshold to justify what not to fund.	☒ major ☐ minor ☐ linguistic	In this context, we refer to willingness to pay as a measure of value of outcomes and not the willingness-to-pay threshold for a QALY. This has now been clarified.
160	39 - 43	1314 - 1524	Section 2.3.8 discusses the methods that can/should be used to estimate QALYs. Perhaps more discussion should be granted on mapping different scales (eg dementia-specific questionnaires back into EQ-5D), for instance following NICE's own guidelines. Perhaps, it would be helpful to suggest research/applied work in terms of mixing information coming from patients and the general public (in a formal way, e.g. through a multilevel model?) to determine the values of QALYs.	☐ major ☒ minor ☐ linguistic	This is partly covered in 2.3.8.2. However, more detailed information is not available in most of the guidelines and is therefore beyond the scope of this project. However, we agree that it would be of value to do a more in-depth analysis of this issue.
161	40	1365	This needs re-wording since it suggest that methods are easier to obtain but it means that data gathered through indirect methods	☐ major	This has been revised to: "The Spanish recommendations specify that QoL weights

			are easier to obtain	☐ minor ☒ linguistic	gathered from indirect methods are recommended since these are easier to obtain, compare and interpret."
162	41	1412 - 1416	– In our opinion for this document to be considered as a guideline EUnetHTA should state an opinion on methods for derivation of QoL weights. – The guidance should be caveated / acknowledged that there are situations when the EQ-5D or other generic preference-based instruments are not suitable (e.g. Autism, Down's syndrome etc.). – It should be acknowledged by EUnetHTA that some populations will not be able to assess their own QoL and proxies may be needed.	☒ major ☐ minor ☐ linguistic	It is out of the scope of this document to recommend methods that not all member organisations can agree on.
163	41	1414	The use of the "QoL" throughout the text should be reconsidered, almost invariably in health-economic evaluation the correct term is stated to be "HRQoL". E.g., "derivation of QoL weights" should be replaced by "derivation of HRQoL weights" If the above is accepted, the definition of "QoL" could be removed from the list of abbreviations	☐ major ☐ minor ☒ linguistic	Even though QoL is used by many of the guidelines, we agree that HRQoL is a better term. We have changed this when relevant.
164	41	1415	"HUI or SF-6D" could be replaced by "HUI, SF-6D or 15D"	☐ major ☒ minor ☐ linguistic	This has been revised accordingly.
165	41	1418	It could be useful in this section to add a statement on quality of life valuation in children as it poses several challenges and may need to rely on the use of proxies. It also poses the question of	☐ major ☒ minor	You are right that the question of using proxies is closely related to the discussion of whose preferences the QALY should represent. However, we have not extracted any information

			accounting for the loss of utility for their parents/caregivers who may be highly impacted.	☐ linguistic	about this from the guidelines.
166	41	1430	I understood that Scotland and England did not both want general public health preferences?	☐ major ☒ minor ☐ linguistic	We confirm that both Scotland and England recommend preferences from the general public.
167	42	1459	It is not clear what the word „consistent" refers to – consistent across countries?	☐ major ☐ minor ☒ linguistic	We have revised to "consistent across technologies and time".
168	42	1460 - 1463	– In our opinion for this document to be considered as a guideline EUnetHTA should state an opinion on by whom the QoL weights should be valued. This paragraph provides no explicit recommendation.	☒ major ☐ minor ☐ linguistic	It is out of the scope of this document to recommend methods that not all member organisations can agree on.
169	42	1474	This should read „data are" not „data is" – data are plural	☐ major ☐ minor ☒ linguistic	We have revised this accordingly
170	43	1485	– *Add* ")"	☐ major ☐ minor ☒ linguistic	We have revised this accordingly.
171	43	1492 -	– In our opinion for this document to be considered as a guideline EUnetHTA should state an opinion on use of	☒ major	It is out of the scope of this document to recommend methods that not all member

		1496	mapping from disease-specific QoL measures to QoL weights that can be used for calculation of QALYs.	☐ minor ☐ linguistic	organisations can agree on.
172	42	1461	The number of differences are not accurately captured by the conclusion. In fact, the differences in how QoL should be captured are a major stumbling block.	☒ major ☐ minor ☐ linguistic	We have tried to make this clearer.
173	44	1540	To avoid confusion „3-5" should be replaced by „3 to 5"	☐ major ☐ minor ☒ linguistic	This has been revised accordingly.
174	44	1543	It is not clear if the recommendation is that costs and outcomes are discounted at the same rate in the base case analysis	☐ major ☒ minor ☐ linguistic	We have added a recommendation concerning that costs and outcomes should be discounted at the same rate in the base case analysis. Since not al guidelines agree with this, we have also recommended to perform several sensitivity analyses.
175	44	1546	"It is also recommended to investigate the effect of reducing the discount rate for health effects and setting both discount rates to zero." could be replaced by "Investigating the effect of reducing the discount rate for health effects and of setting both discount rates to zero is also recommended."	☐ major ☐ minor ☒ linguistic	This has been revised accordingly.
176	44	1553	No country specific guidelines would talk about converting costs to relevant currencies since the expectation is that you use costs specific to that country. If this point is relevant to the very	☐ major ☒ minor	We understand your point. However, some guidelines contain information about how to do this.

			specific case when you develop a model for Europe (which does not often happen) then it should be clearer. This is also inconsistent with earlier text which talked about using national sources for unit costs	☐ linguistic	Economic evaluations conducted within the HTA core model are conducted with the aim of informing decision makers in Europe and not only one single country. However, we have decided to remove the recommendation of using euros within EUnetHTA since this is not based on the information in the guidelines.
177	44	1156	CatSalut also prefers the reference year to be the current year or the most recent one.	☐ major ☒ minor ☐ linguistic	CatSalut has been included in this list now.
178	45 - 46	1570 - 1655	We believe that it may be helpful to be even more prescriptive, suggesting that a set of graphical displays should always be provided. After all, computer packages exist that can standardise the production of such tools.	☐ major ☒ minor ☐ linguistic	Thank you for this good advice. However, it is not based on the guidelines from the member organisations and can therefore not be included here. Nevertheless, it is a good advice for future guidelines.
179	45	1594	It is noted that price indices should be used to update cost figures, but each country is listing its own preferred, national index. Differences between indices make for limited comparability of the analyses. Also, the recommendation to use Euros as a currency unit seems to go counter to several guidelines – and may not be accepted in several additional countries.	☒ major ☐ minor ☐ linguistic	Economic evaluations conducted within the HTA core model are conducted with the aim of informing decision makers in Europe and not only one single country. However, we have decided to remove the recommendation of using euros within EUnetHTA since this is not based on the information in the guidelines.
180	45	1597	Does the phrase „within EUnetHTA" mean across countries or in individual countries since if it's the latter, then this is not relevant for some countries (eg England and Scotland will not	☐ major ☒ minor	Economic evaluation conducted within the HTA core model are conducted with the aim of informing decision makers in Europe and not only one single country. However, we have decided to

			generate analyses in euros for submissions to England and Scotland)	☐ linguistic	remove the recommendation of using euros within EUnetHTA since this is not based on the information in the guidelines.
181	45	1598	– We suggest removing the recommendation to express economic evaluations in Euro. – In our opinion this does not make sense for dossiers to assume the perspective of another member state which might not use the Euro a. The unit costs of many itemized inputs as recommended in the costing section (2.3.6.2) would give a false sense of transferability of results	☒ major ☐ minor ☐ linguistic	Economic evaluation conducted within the HTA core model are conducted with the aim of informing decision makers in Europe and not only one single country. However, we have decided to remove the recommendation of using euros within EUnetHTA since this is not based on the information in the guidelines.
182	46	1624 - 1625	Do you mean 20K to 30K £ in comparison to a different health technology with a different range?	☒ major ☐ minor ☐ linguistic	We are sorry but we do not understand this comment. The way it is phrased in our text is exactly the way it is phrased in the NICE guidelines. It means that a QALY gained is equivalent to a gain of 20 000 or 30 OOO £.
183	46	1641 - 1655	Please give citations in which all these desired requirements have been implemented. If a substantial number of publications can not be found nor can a rationale be found to bring all this information in a feasible way into an acceptable publication, then these desired requirements should be simplified.	☒ major ☐ minor ☐ linguistic	We are sorry but we do not understand this comment.
184	46	1642	This should be rewritten for clarity for example „results should be presented in a simple, disaggregated form"	☐ major ☐ minor ☒ linguistic	This text comes from the Economic Evaluations (ECO) domain of the HTA Core Model. We have revised this slightly and have also forwarded the information to WP8.
185	46	1647 - 1655	– In our opinion the value of this document will be increased if EUnetHTA can provide a sample format for how results should be presented (e.g. shell-table similar to NICE guidelines). We feel that this would increase the consistency	☒ major ☐ minor	It is out of the scope of this document to recommend methods that not all member organisations can agree on. However, we agree and will recommend this for future

			of reported results and usefulness for those less familiar with economic evaluations.	☐ linguistic	collaborations.
186	46	1650 - 1652	This needs rewritten – it talks about incremental then absolute effects then talks about separate then incremental effects again	☐ major ☐ minor ☒ linguistic	We have tried to make this section clearer.
187	47	1672	This section suggests that ICERs should be reported with confidence intervals. We argue that these are actually irrelevant. When computing an ICER, uncertainty in both individual level variation and parameters are being averaged out. Thus, the ICER describes the decision given *current* knowledge, in which uncertainty really does not play a role. Uncertainty does play a role, of course, because current evidence may not be conclusive and thus there is the option of delaying the final decision until new data are collected – that's the rationale of PSA. But the points populating the C/E plane are not ICERs - the ICER is the central point of that distribution. We suggest that both the ICER and C/E plane should be provided, though. (Note that section 1.1.1.4 makes a similar confusion by referring to the distribution of the ICERs; this should read as 'joint distribution of cost and benefit differentials').	☒ major ☐ minor ☐ linguistic	In this section we discuss what the different guidelines say.
188	47	1672	This section suggests that ICERs should be reported with confidence intervals. We argue that these are actually irrelevant. When computing an ICER, uncertainty in both individual level variation and parameters are being averaged out. Thus, the ICER describes the decision given *current* knowledge, in which uncertainty really does not play a role. Uncertainty does play a	☒ major ☐ minor ☐ linguistic	Duplicate, see 187

			role, of course, because current evidence may not be conclusive and thus there is the option of delaying the final decision until new data are collected – that's the rationale of PSA. But the points populating the C/E plane are not ICERs - the ICER is the central point of that distribution. We suggest that both the ICER and C/E plane should be provided, though. (Note that section 1.1.1.4 makes a similar confusion by referring to the distribution of the ICERs; this should read as 'joint distribution of cost and benefit differentials'). Baio, G. (2012). Bayesian Methods in Health Economics. Boca Raton, FL: Chapman Hall, CRC Jackson, C. H., Sharples, L. D. & Thompson, S. G. (2010). Structural and parameter uncertainty in Bayesian cost-effectiveness models. Journal of the Royal Statistical Society, Series C 59: (2), 233-253 Jackson, C., Bojke, L., Thompson, S. G., Claxton, K. & Sharples, L. D. (2011). A framework for addressing structural uncertainty in decision models. Medical Decision Making 31: (4), 662-674		
189	47	1679	Before conducting "model meta-analyses" potential sources of heterogeneity should be explored. Model averaging needs to be justified.	☒ major ☐ minor ☐ linguistic	No change is done since the text only presents what the French guidelines recommend.
190	47	1694	We welcome the fact that the guideline mentions the EVPI. It is probably a reflection of the state of the art, which is relatively limited in terms of application of this tool, that this mention is	☐ major ☒ minor	This document is based on the extractions of information of the guidelines used by EUnetHTA-members.

			only minor. However, there are currently several applied and methodological developments surrounding the use and availability of the EVPI and, particularly, the EVPPI (for example, R packages as well as web applications that allow the easy computation of these quantities for general health economic models). These should be encouraged and advertised more thoroughly – outlets such as this guideline document are among the best ways of doing this.	☐ linguistic	
191	48	1704 - 1708	– If EUnetHTA are recommending a PSA, we feel it would be beneficial to state their opinion to graphically display results (e.g. scatter plot of cost-effectiveness-plane, cost-effectiveness acceptability curve).	☒ major ☐ minor ☐ linguistic	We agree that this would be useful. However, this document is based on the extractions of information of the guidelines used by EUnetHTA-members and the different guidelines seldom present that detailed information.
192	48	1708	– It would be valuable to know EUnetHTA's opinion on the technique for displaying uncertainty (e.g. particular modeling techniques for PSA)	☐ major ☒ minor ☐ linguistic	We agree that this would be useful. However, this document is based on the extractions of information of the guidelines used by EUnetHTA-members and the different guidelines seldom present that detailed information.
193	50-51	1795 - 1797	The different views can also be a reflection of different values of HT assessors, appraiser and nations.	☒ major ☐ minor ☐ linguistic	We believe that this is covered by "different contexts and policies" but we have added some text to make it clearer.
194	51	1798 - 1801	End-users of HTA information should rather be involved in the beginning of any EUnetHTA project. Any future EUnetHTA project should need to be founded on a business case from the perspective of the end user of HTA information.	☒ major ☐ minor ☐ linguistic	We have forwarded this comment to the coordinators.
195	49-	1749	The document is presented as "methodological guidelines" and	x **major**	The guideline aims to develop a general

	50	- 1756	not as a review of existing practices. This imposes to review the robustness of methods because proposing any recommendations. The European project ECHOUTCOME,which has been referenced in the document, has established that the QALY outcome is flawed, based on a European experiment testing each underlying assumptions. Why the document "recommend" to present outcomes im QALY ? because some European countries use it ? Firstly, member states using QALY as reference cases are not the majority in Europe. Secondly, whatever the majority is, sciences is not majority ! it is scientific demonstration. It is not a question to say that the "use of QALY is debated", but to say that European research has established that the QALY approach is flawed, leading to other European guidelines (Echoutcome European Guidelines) to abandon QALY in HTA.	☐ minor ☐ linguistic	framework for HTA-reports conducted within EUnetHTA. Since these projects aim to inform decision makers in the different European countries, we would argue that it is important that EUnetHTA considers what the decision makers in the European countries are asking for when recommending methods for economic evaluation. In all but 4 of the 25 guidelines, QALYs are recommended as one of the main outcome measures. We have added some text in relation to the text about that the use of QALYs has been debated in the health economic literature.
196	58	2103	The reference should be the following: Puig-Junoy J, Oliva-Moreno J, Trapero-Bertrán M, Abellán-Perpiñán JM, Brosa-Riestra M y **Servei Català de la Salut (CatSalut)**. Guía y recomendaciones para la realización y presentación de evaluaciones económicas y análisis de impacto presupuestario de medicamentos en el ámbito del CatSalut. Versión 1.0. Generalitat de Catalunya. Departament de Salut. Servei Català de la Salut: Barcelona, Octubre 2014.	☐ major ☒ minor ☐ linguistic	This has been revised accordingly.
197	60-64	2125	Please add how often these guidelines have been applied and used for decision making per type of technology.	☒ major ☐ minor	This is not in the scope of the current project.

				☐ linguistic	
198	66	2138	Regarding CatSalut (Catalonia-Spain), economic evaluations are not only to supposed to provide information but also to provide guidance to the decision making process when issuing recommendations on the use of a specific medicine within the Catalan Health Service.	☐ major ☒ minor ☐ linguistic	This has been revised accordingly.
199	80	2185	CatSalut also accepts meta-analysis and indirect comparisons when there are not direct comparative studies.	☐ major ☒ minor ☐ linguistic	This has been revised accordingly.
200	82	2194	QALY has never been a prefered measure in countries such as France, Italy, Russia, etc ! Reimbursement authorities do not require QALY as reference case in these countries. For example very few QALY studies have been performed in France. What is the source of this false statement ?	x **major** ☐ minor ☐ linguistic	This guideline is based on the recommendations in national or regional guidelines and we believe that the extractions from the guidelines are correct. For example, the guideline by HAS clearly states "If the main objective of the intervention is improving life expectancy, the main health outcome measure is the QALY. However, if the main objective of the intervention does not imply an effect on quality of life, HAS recommends a CEA with costs per life-years gained as the outcome measure instead of a CUA."
201	84	2208	CatSalut also accepts WTP methodology in CBA.	☐ major ☒ minor ☐ linguistic	This has been revised accordingly
202	89	2233	In the CatSalut's guidance, costs should be also adjusted to the reference year.	☐ major	This has been revised accordingly

				☐ minor ☐ linguistic	

Participants

Color in table	Name of organization / company / project	SAG Member	Name of contact person / person who submitted comments
	KCE – BE- Belgian Health Care Knowledge Centre		Frank Hulstaert, MD, MSc, FBCPM
	Sanofi-Pasteur - MSD		Vanessa Remy
	Novo Nordisk	X	Mark Aagren
	EFPIA	X	Karianne Johansen
	EDMA	X	Karsten Berndt Chair of EDMA HTA Task Force
	ECHOUTCOME – FP7		Ariel Beresniak, MD, MPH, PhD Echoutcome (Fp7) Project leader
	Genzyme – Sanofi		Alicia Granados MD, PhD PH
	ESMO - European Society for Medical Oncology	X	Rosa Giuliani
	Mapi Ltd		Madeleine Steer
	F. Hoffmann-La Roche Ltd, Division Pharma		Joshua Ray , Group Head Health Economic Modelling
	PSI – HTA-SIG (Statistician of the Pharmaceutical Industry, HTA Special Interest		Patrick Moneuse - Global Head of

	Group)		Biometrics, Vifor Pharma
	THL – FIN - National Institute for Health and Welfare		Neill Booth
	CatSalut – ES - Servei Català de la Salut		Ruth Puig-Peiró, MSc