ISPOR
The professional society for health economics and outcomes research

Improving healthcare decisions

505 LAWRENCE SQUARE BLVD SOUTH
LAWRENCEVILLE, NJ 08648

P +1-609-586-4981
F +1-609-586-4982

info@ispor.org

November 1, 2022

Dear EUnetHTA:

ISPOR – the professional society for health economics and outcomes research - is pleased to respond on behalf of its membership to your Methodological Guidelines consultation, **"D4.4 Endpoints."** We thank you for the opportunity to comment on these draft guidelines.

ISPOR is a scientific and educational society with many of its members engaged in evaluating health technologies, including pharmaceuticals, medical devices, and other interventions. We have a large membership living and working in 110 countries globally; nearly 20% (1 in 5) of our membership resides within the European Union. Members across our organization come from a range of disciplines, including health economics, epidemiology, public health, pharmaceutical administration, psychology, statistics, medicine, and more, from a variety of stakeholder perspectives, such as the life sciences industry, academia, research organizations, payers, patient groups, government, and health technology assessment bodies. The research and educational offerings presented at our conferences and journals are relevant to many of the issues and questions raised in this request for information.

The response to this consultation was led by members of our Clinical Outcomes Assessment (COA) Special Interest Group, with comments solicited from several of our membership groups, including our the Health Science Policy Council, HTA Roundtables, Institutional Council, Rare Disease Special Interest Group, Patient-Centered Special Interest Group, Statistical Methods in HEOR Special Interest Group, Oncology Special Interest Group, and authors of previous ISPOR Good Practice Reports about COAs and patient-reported outcomes (PROs). The attached document provides a synthesis of their comments, summarized by an expert panel and our Chief Science Officer. We hope they prove useful.

ISPOR would be happy to answer any questions about our response, as well as participate in any follow-up consultations on the relevant program items mentioned within the report.

Sincerely,

Nancy S. Berg
CEO & Executive Director
ISPOR

EUnetHTA 21 Public Consultation

Comments should be submitted not *later than* *30 August 2022, 23:59 CET*

D4.3.1 Practical Guideline Direct and Indirect Comparisons,
D5.2 JCA Assessment Report Template,
D7.2/3 Guidance and template for the interaction with patient representative, healthcare professional and other experts ***(please note this consists of four templates)***

Please use this form for submitting your comments and share your completed comment form to JCA_Secretariat@zinl.nl prior to the deadline (1 November 2022, 23:59 CET). When submitting your comment form, please include "**EUnetHTA 21 – Public Consultation – D4.4** in the subject line of your e-mail.
Please carefully read the principles for public consultation here, prior to your review, as these are binding for our process.

We kindly ask you to:

1. Submit one consolidated response per organisation; in a word-file
    a. PDF files will not be accepted;
2. Complete the first table; if this table is not completed, the input will not be considered by EUnetHTA 21;
3. Put each new comment in a new row;
    a. Please be clear about the context of your comment and if possible, provide a suggestion for rewording;
    b. Please consider the HTA Regulation (EU) 2021/2282 when reviewing the document and when you provide comments;
    c. Please consider the corresponding project plan when commenting. Comments that refer to matters out of the scope of the deliverable may not be considered by EUnetHTA 21.
    d. Please do not provide linguistic comments, as the document will undergo language editing prior to finalization;
4. Insert the page number and line/section number on which your comment applies. If your comment relates to the document as a whole, please put **'general'** in this column;
5. Provide a description of your comment as specific as possible and preferably also provide a suggestion for rewording. If you wish to draw our attention to published literature, please supply the full reference;
6. Add rows as needed.

NB: All comments received within the deadline of the consultation and following the correct format will be published on the website, together with the final deliverable. Only comments eligible for consideration will be answered by EUnetHTA 21. The answers will be made publicly available as well. EUnetHTA 21 may decide to rank the comments received on importance.

Please complete this table. If this is not completed, your comments will not be considered.

Name organisation & abbreviation	ISPOR – The Professional Society for Health Economics and Outcomes Research
Country	Headquarters is based in the USA, but nearly 20% (1 in 5) of our membership lies within the European Union.
Contact details (name & e-mail address) – *this information will not be published*	Kelly Lenahan, Associate Director, Content Strategy and HTA – klenahan@ispor.org Richard Willke, Chief Science Officer – rwillke@ispor.org We would also like to thank Jessica Roydhouse, PhD, BA, MPH (Hons), Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS, Australia and Jagadeswara Rao, MBA, PharmD, PhD, Center for Observational & Real-world Evidence (CORE), Merck, Kenilworth, NJ, USA for their help with consolidating all the comments received from ISPOR members.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Insert your name and organisation ***Please repeat in each row***	***Insert 'general' if it relates to the whole document*** ***Please don't put 'p' before the number***		***Please insert each new comment in a new row.***	***Please indicate with 'x' if your comment is an editorial comment.***
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	General		It is suggested to look beyond the references for standardization purposes. The definitions could be a bit more clear by using some more up to date documents, and citation of regulatory documents or papers from regulatory bodies (e.g., EMA reflection papers, FDA guidance) would support harmonization across the sector. Further to this point, and to the goal of harmonization and standardization, it would also be helpful if definitions for key terms, such as those around COAs, came from recently published documents or drew upon commonly used/widely cited documents from professional societies. One very relevant document is	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
				the recent FDA draft guidance on Clinical Outcome Assessments: https://www.fda.gov/about-fda/clinical-outcome-assessment-coa-frequently-asked-questions	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	General		The guideline in its current version summarizes general concepts but still leaves large room for interpretation and specific application of concepts between the member states (MS): Based on scientific rationale there should be more harmonization between MS regarding the following aspects: - Clinical relevance - Acceptance of surrogate endpoints - Acceptance of Responder definitions	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	General		Much of the guidance is stated in fairly general terms and refers to good clinical practice or good statistical practice, which is per se meaningful as some decisions have to be made on a case to-case basis. However, a key concept of the entire JCA is also to have sufficient certainty about the acceptance of the approaches used by the HTD and on the implementation of good scientific practice in the specific situation of interest. Therefore, the key parameters for implementing outcomes, surrogate validation, validity, reliability and responder definitions should be determined in advance by a close exchange between assessors and HTD – on the basis of state-of-the-art scientific methods. The outcome of this exchange should be binding in the sense that the agreed methods are accepted in the final assessment	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	General		The guideline should be open for innovative methods established after this guideline comes into effect. A corresponding review process for an update of the guidelines should be implemented to ensure that the guideline reflects the current state-of-the art.	
D4.4 Endpoints	ISPOR – The Professional	General		It would be helpful to have a process map clarifying roles of EMA, EUnetHTA, HTA, HTD, and MS in defining,	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Society for Health Economics and Outcomes Research			requesting, submitting, and assessing COA in the JCA at the beginning.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	General		Since there seems to be some variation in volume among the items, we suggest that the existing guidelines be used as a reference and that a little more balance be achieved.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	General		The document could be more clearly divided into recommendations for Member States and recommendations for how to write the JCA report. We would also welcome an overall aims description for assessing the outcomes and impact of pharmaceutical products. Pharmaceutical products are not only with the aim to prolong life, but also for patients to be healthy and productive, to enable work participation, to enable living in own home etc. Especially given the demographic challenges we are facing in most countries. To keep chasing mortality as main outcome seems outdated to me. I do not think this guidance is reflecting the needs in the health care sector and in societies today. In addition, the use of proxies/surrogate endpoints should be encouraged where these can be used to (adequately) predict other desired outcomes. This should be so for ethical and economic reasons. I think the document should consider the totality of the health care environment in the future.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	4	73	Missing acronyms that could be interesting to add and to mention in the text: - PRO-CTCAE, as FDA and EMA acknowledge the value of self-report of symptomatic AE by patients to complement the physician report of AE - apart of the PGRC, another global assessment is the PGIS: patient global impression of severity; COA – clinical outcome assessment. EQ-5D is missing	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	5	1.1 Problem statement, scope and objectives (Lines 75-116)	It isn't clear if the guidance is proposing a set of harmonized recommendations that all member states need to follow or if the guidance is allowing member states to select certain recommendations that are more applicable and relevant to them. It is critical to strike a balance in proposing harmonized recommendations and accommodating the varying requirements of each of the member states at national level. For JCA, aspects pertaining to measurement scales / instruments for assessing outcomes including PROs need to be harmonized across member states.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	5	85-88	The judgement of what constitutes a clinically relevant outcome, responder definition, etc. should be harmonized across member states (MS) based on a scientific rationale. Please consider changing the sentence as follows: "While MS are required to give due consideration to the JCA reports published (Article 13 (1)), the rating of the additional benefit of a treatment may differ at a national level which is based on the clinical relevance of the measure of relative effectiveness.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and	5	91	Does this refer to rank ordering of endpoints? If so - this is contradictory to what regulatory agencies usually require (rank ordering and prespecifying as a prerequisite to granting labeling language around those endpoints). Also in reference to ranking on health outcomes,	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Outcomes Research			including weighting and relevance of primary vs. secondary vs. exploratory endpoints: Often, payer-related endpoints are secondary or exploratory given the trial is not powered to show differences. How will these endpoints be considered in the evaluations? In addition to the JCA assessments, will member states require or ask for data for patients in their countries? What if there are no patients in the trials for that country?	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	5	96-99	The guideline says that "the conclusions that MS can draw regarding the clinical added value of a treatment can be impacted by factors such as appraisal of the validity and reliability of the measurement scales of instruments or of the relevance of intermediate or surrogate outcomes." The assessment of the validity and reliability of the measurement scales of instruments should be based on scientific standards and should be harmonized across member states.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	5	100-105	This guideline can also be useful for the submitters.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	5-6	1.2 Relevant articles in Regulation (EU) 2021/2282 (Lines 117-123)	Providing links to these articles would be helpful.	
D4.4 Endpoints	ISPOR – The Professional Society for Health	6-7	2.1 Definitions (Lines 124-177)	It would be helpful to mention the term clinical outcome assessment (COA) and align with existing and commonly used definitions of specific COAs such as patient-reported outcomes (PROs), clinician-reported outcomes	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Economics and Outcomes Research			(ClinROs), performance outcomes (PerfOs), proxy-reported outcomes (ProxROs) and others. In addition, it is worth nothing that there have been professional society task forces and documents that have discussed COAs such as ClinROs, PerfOs, and ProxROs. ISPOR Task Forces have produced reports to help standardise ClinROs and PerfOs and ISOQOL has a Task Force on proxy reporting/ProxROs with a similar objective. Furthermore, it would be worthwhile to note existing regulatory definitions around COA terms such as those used by FDA and EMA. Although it is certainly relevant to note the growth and importance of digital data, it may be worth making a distinction between digital data and PerfOs; as currently written, the paragraph could be interpreted as suggesting an inherent link between the two. In addition, the proposed definition of proxies is relatively non-standard, and it may be beneficial to cite regulatory or other existing definitions of proxies. Furthermore, although carers/caregivers can provide proxy reports, in some situations the carer/caregiver experience on its own may be relevant, and thus caregiver outcomes could be mentioned. Distinguishing between proxy-reported outcomes and observer-reported outcomes would also be helpful.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	6	140-143	It seems “absolute effect” is one type of difference measure. If so, we may reword it as below. “… effect measures are either difference measures (e.g., absolute effect, mean difference in change, risk difference) or ratio measures (e.g., risk ratio, odds ratio, hazard ratio). However, other statistics can be used to express other aspects of a treatment effect such as the absolute effect or a within-group change.”	
D4.4 Endpoints	ISPOR – The Professional	7	176	In addition to DAS, the Mayo Score is another hybrid measure that can be added. The Mayo Score was	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Society for Health Economics and Outcomes Research			developed as a composite disease activity index for use in clinical trials. The original description of the Mayo Score included an assessment of 2 patient-reported outcomes [PROs; stool frequency (SF) and rectal bleeding (RB)], the endoscopic appearance of the mucosa (endoscopic score, ES), and a Physician's Global Assessment (PGA), each of which were scored on a scale from 0 to 3, giving a maximum total score of 12. Reference: Schroeder KW, Tremaine WJ, Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. N Engl J Med.1987; 317:1625–1629.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	6	126-127	Suggestion for lines 126 and 127: Health outcome is the impact that a specific health intervention, technology, policy, or program has on a person, group, or population. This endpoint can be any changes in morbidity (efficacy, effectiveness, and safety) and/or mortality. Safety and effectiveness impact on HRQOL and QOL.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	6	138-139	138 and 139: instead of existing 'health technologies', we can say 'health interventions' because the comparator can be a surgical method or a public health intervention, not a health technology.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	6	147-151	Suggest adding the acronym here (ClinRO). The term "clinically reported outcomes" is inconsistent with the sentence before it (Clinician-reported outcomes), which is the usual term for these measures. We encourage use of the Clinician-reported outcomes, or ClinRO, term, and suggest these measures be referred to in a consistent manner.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	6	152	Technology assessed outcome measurements could be made through digital solutions, diagnostic tools (cont glucose measurements), AI and algorithms. To some extent they are mentioned in a section further into the document, however, they could also be mentioned under technology assessed outcomes	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	6	161-166	It might be better to distinguish “observer-reported outcome” and “proxy-reported outcome” here. “...reported by an observer with shared experience. An example would be a caregiver if the patient is unable to answer the items. These cases are referred to as PROs answered by "proxies” ….”	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	6	164	Can there be recognition of Caregiver reported outcomes, as a standalone measure and resulting from the intent to measure the caregiver perspective (i.e separate from their role as a proxy?). This will allow assessment of the broader impact of a condition beyond individual patients’ experiences	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	7-8	2.2 Summary	It may be beneficial to extend the recommendation regarding request formulation to mention international standards such as SISAQOL (standardisation of PRO analyses in cancer trials). Furthermore, rephrasing the recommendation to suggest that good clinical and statistical practice be incorporated would help to remind readers of the importance of these issues. Additionally, it would be worth mentioning the importance of pre-specifying statistical methods as part of these general recommendations.	
D4.4 Endpoints	ISPOR – The Professional Society for	7-8	2.2 General considerations (Lines 178-208)	In this part or somewhere, is there any definition of difference between preference-based measures (PBM) and non-PBM (including COS) or utility?	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Health Economics and Outcomes Research				
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	7-8	178-208	It would be good to consider and clearly state RWE outcomes as well as the outcomes from RCTs in these three paragraphs.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	7-8	178-210	The wording for the scoping process on the outcomes should be accompanied by justification/rationale, in keeping with good statistical practice, to enable HTD to better understand the request. These should also be aligned to the study design, otherwise it becomes an immediate disadvantage to the HTD for not being able to provide them.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	7	183-184	The first sentence of the paragraph is unclear, mainly with regard to whether it refers to the MS or the HTD. Also, sometimes results are not best 'obtained‘ from broader concepts, but rather are better 'obtained' from specific items or subscales of HRQoL.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	7	190	Does "assessed differently" mean using a different instrument? Please be more specific.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	7	191-192	General considerations refer to the definition of an outcome, in particular specifying an appropriate PRO endpoint. The guideline says, “To alleviate this issue, a general recommendation could be to formulate a request as such: “[Outcome of interest] measured preferably as [insert measure]”.” The recommendation should be extended, and reference given to the SISAQOL-IMI2 initiative aiming for setting International Standards in Analysing Patient-Reported Outcomes and Quality of Life Endpoints in Cancer Clinical Trials. Similarly, to the reference to COMET for COS funded also by IMI2. Thus, please consider adding “To alleviate this issue, a general recommendation could be to formulate a request as such: “[Outcome of interest] measured preferably as [insert measure]” and to consider international standards in analyzing PROs and Quality of Life Endpoints in cancer clinical trials as provided by SISAQOL-IMI.”	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	7	192-201	“Timing of outcome assessment” should be justified in the context of the kind of trial and the treatment and follow Good Clinical and Statistical Practice. For example, it could be inappropriate to ask for an OS delta in neoadjuvant treatment at 12 months because other endpoints are more appropriate.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	7	195-196	There should be recognition that with PROs there are limitations in how long they can be assessed within a study. Also, general flexibility for HTDs in designing the schedule of assessment in their studies would be appreciated, as treatments may have different onset of action. The schedule of assessments should reflect that difference.	
D4.4 Endpoints	ISPOR – The Professional Society for	7	195-198	The guidance addressed concerns if the follow-up was considered not sufficiently long in the clinical study submitted as evidence and recommended to formulate a	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Health Economics and Outcomes Research			request “[Outcome of interest] measured preferably at [insert timing of assessment]”. Preferences may vary between MS. One MS might be interested in ““rate of major adverse cardiovascular events 2 years after inclusions” another MS is interested in 12 months or 18 months, 3 years after inclusion. On the other hand, the data collection of a study usually covers the clinically most important periods, where changes in symptoms or side effects are expected. In addition, data collection period and frequency of assessments also consider the factors that would compromise a scientifically sound evaluation e.g. when the rate of missingness is considered too high to draw reasonable conclusions and also acceptable level of patient burden. In essence the request for a timing of assessment should not only be harmonized across member states, it should also adequately reflect the disease setting and clinical context. So, a joint scientific advice meeting (including REG and HTA) is recommended to clarify the needs so that the study can be designed accordingly. Second, as marketing authorization could be based on a positive benefit-risk evaluation using interim results of a study, only sparse data might be available to provide a reasonable precise estimate e.g., for the rate of major adverse cardiovascular event 2 years after inclusion. Modelling approaches may be informative to estimate a specific outcome at a certain time point accordingly. Such methods could be useful to decrease uncertainty and thus should be taken into account for the added benefit assessment. In summary, please consider changing the wording from: “A general recommendation could also be to formulate a request as such: “[Outcome of interest] measured preferably at [insert timing of assessment]”. To “A general recommendation could be to align key time points of interest prior to study initiation in a joint scientific advice meeting, or to pre-specify statistical methods to	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
				estimate the effect at certain time points also using modeling approaches. “	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8	2 Summary Table (Line 209)	With reference to the following sentence “Effect measures should not be specified by MS. The HTD is responsible for presenting results using appropriate effect measures in accordance with good clinical and statistical practice”. Please clarify the criteria/guidelines that the HTD should consider when selecting effect measures.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8	2 Summary Table (Line 209)	Bullet 1: Should “or safety” be added to the end of this sentence?	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8	2 Summary Table (Line 209)	General comment to the requirement for member states in the assessment scoping process – please include a request for the rationale underlying their requests for specific outcomes and measures.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8	2 Summary Table (Line 209)	The last bullet is not worded clearly. We suggest rewording as follows: “An accurate definition of any reported outcome is required and would include a description of the concept, source of information, the measure of the outcome, timing, and effect measure.”	
D4.4 Endpoints	ISPOR – The Professional Society for	8-9	3.1 Definition of patient-centred outcomes (Lines	It would be helpful to differentiate patient-centered care with physician centered care using examples.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Health Economics and Outcomes Research		210-250, inclusive of box)	In addition, it would be helpful to suggest the 'essential' vs 'desirable' outcomes needed for JCA to HTDs. While OS is relevant, it may be infeasible for HTDs to measure OS in early-stage cancers. It is equally important not to ignore patient-relevant surrogate outcomes. Recommending the use of validated tools only might pose challenges in certain therapeutic areas where such validated measures are not available. It also limits the innovative use of endpoints or measures from RWD sources, even if only as exploratory endpoints.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8-9	3.1 Definition of patient-centred outcomes (Lines 210-250, inclusive of box)	What is the difference between the "patient-centered outcomes" presented here and the "patient-relevant outcomes" presented in previous guidelines? Also, can "patient-centered outcomes" be considered to include PROs?	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8-9	223-224	To understand the "contrast" meant here, it would be helpful to define physician-centred care. Do you mean clinician-reported outcomes, or physician services as typically captured in quality of care measures, or something else? Mortality is clearly important to patients, but is generally clinician-reported. Thus understanding what is meant by physician-centered care will help clarify the contrast with, and meaning of, patient-centered outcomes.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8	212-220	We acknowledge that outcomes supporting the benefit-risk assessment might be "less suitable for the needs of JCA". However, the healthcare and treatment decisions strongly depend on the prescribing information (PI). Hence all outcomes described in the PI should be considered for JCA purposes and used as a common "core" outcome set among all member states. Please consider changing the sentence to "Some outcomes may be fully acceptable as support for the risk/benefit ratio	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
				assessment of a certain therapy but are less suitable for the needs of JCA. However, outcomes described in the prescription information should be considered for JCA purposes." "The acceptability of an outcome is subject to MS interpretation of their relevance within their national process for decision making and thus may differ between MS. At least outcomes described in the prescription information should be considered as a core set of outcomes. "	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8	215-216	How does this apply to an acute condition, e.g., infection?	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8 & 9	217 & 237	Death and mortality are highlighted as outcomes, but many diseases are not fatal or even have "irreversible" events. Some mention of other important outcomes, eg, pain relief, mental health, and return to work seems warranted.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	8	224	Anchoring on the emphasis on patient-centered outcomes as those which matter to patients, one should try to bring in the relevance and importance of surrogate outcomes (not withstanding what has been said above in line 213-215 about surrogate outcomes) because of the importance that patients might ascribe to those outcomes in early-stage disease (i.e., oncology).	
D4.4 Endpoints	ISPOR – The Professional Society for	9	236	Please provide a more inclusive definition of long-term outcomes, since not all conditions are chronic or fatal.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Health Economics and Outcomes Research				
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	9	236-242	Marketing authorization of a drug could be based on a positive benefit-risk evaluation using interim results of a study based on dual endpoints one of which demonstrated superiority (e.g. PFS in a cancer study) and the other (overall survival OS) show at least a clear trend. A final OS outcome is rarely available as studies are still ongoing at the time of JCA. Additional endpoints such as PFS2 (time to to second objective disease progression, or death from any cause, whichever first) and time to next anti-cancer treatment could be considered as intermediate endpoints to substantiate the trend observed for an interim OS outcome. Interim analyses may provide only sparse data to provide a reasonable precise estimate for long term survival (e.g., 5-years survival rate). intermediate endpoints and modelling approaches could be useful to decrease uncertainty of long-term outcomes and thus should be taken into account to assess the added value of a health technology. Another aspect concerns maintaining the study integrity of a trial. For instance, if interim results of a randomized controlled double blind clinical study qualify for marketing authorization, independent data monitoring committees may suggest keeping the double-blind nature of the trial to provide an unbiased OS estimate. HTDs may have to decide whether to risk losing a full approval or a positive benefit assessment by MS when unblinding based on strong intermediate outcomes. Early access to an efficacious drug could be compromised by conflicting requirements and thus decisions might not be considered patient centered. Please consider changing the following sentence as follows:	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
				If it is not feasible to measure a final outcome, then intermediate or surrogate outcomes may be acceptable. Acceptability should consider if there is evidence of a strong association or correlation of effects on the surrogate or intermediate outcome with the effect on the final outcome [14], study integrity considerations e.g. when studies need to be unblinded to obtain the final outcome and pre-specified statistical approaches to estimate the long-term using modelling approaches. Scientific advise meetings are recommended to clarify opportunities.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	9	239-243	Please develop a method/matrix table to clearly specify which outcomes will be considered as “essential” for HTA evaluation, depending on the specific disease (e.g. metastatic vs. early-stage cancers). This would help HTDs develop clinical trial protocols to collect outcomes that are meaningful for HTA evaluations.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	9	242	Alternatively (to the suggestion of strong correlation), given the acknowledgment around patient-centered care in the paragraph starting line 223, surrogate outcomes need to be considered acceptable if patients find them meaningful.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	9	248	Would avoid mention of 'validated' in this context in which clinical scales are referred to. A large proportion of such clinical measures aren't validated relative to the rigor of similar efforts expected of PROs. Alternatives: validated tools...or those developed by professional clinical associations and guidelines developers.	
D4.4 Endpoints	ISPOR – The Professional	9	250	In the “Points of attention for the assessment of scoping process” box: Given the acknowledgment around patient-	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Society for Health Economics and Outcomes Research			centered care in the paragraph starting line 223, surrogate outcomes need to be considered acceptable if patients find them meaningful. "If it is not feasible to measure final outcomes, then intermediate or surrogate outcomes may be acceptable if there is evidence of a strong association or correlation of effects on the surrogate or intermediate outcome with the effect on the final outcome." Suggestion: If it is not feasible to measure final outcomes, then intermediate or surrogate outcomes may be acceptable if there is evidence of a strong association or correlation of effects on the surrogate or intermediate outcome with the effect on the final outcome OR evidence that patients find the intermediate and/or surrogate outcomes to be meaningful.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	9	250 (box)	To reiterate, long-term or final outcomes do not seem defined widely enough to cover all conditions.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	9	252-254	It would be most helpful to provide a method/matrix table to clearly specify which outcomes will be considered as "essential" for HTA evaluation, depending on the specific disease (e.g. Metastatic vs. early-stage cancers). This would help HTDs appropriate develop clinical trial protocols to collect outcomes that are meaningful for HTA evaluations.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics	9	259	If "multimorbidity" is meant as a specific situation and not simply as a general example, it would be helpful define a "multi-morbidity" condition in terms of the combination, severity, and number of conditions.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	and Outcomes Research				
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	9	266-272	The recent FDA draft guidance, "Core Patient-Reported Outcomes in Cancer Clinical Trials," would be a most helpful complement to this paragraph: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/core-patient-reported-outcomes-cancer-clinical-trials.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	9	277-280	1. This is the only mention of generic multiattribute utility instruments in the document; more context about their use as endpoints would be helpful. 2. Given their importance in creating QALYs for CEA – even if not JCA – we would encourage their inclusion as complementary endpoints in trials *whenever possible*.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10	281	Cardiovascular diseases are the leading cause of death worldwide: 1. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10	284	Core clinical outcomes sets are hardly universal, and many(most) indications may not have these. Suggest providing guidance on what instrument characteristics should be assessed to understand their validity for the specific treatment and indications (e.g., they are fit for the context of use).	
D4.4 Endpoints	ISPOR – The Professional	10	284	After the sentence "Specific definitions of outcomes typically used in oncology are provided in Appendix A,"	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Society for Health Economics and Outcomes Research			please provide a short list of key outcomes for rare disease ideally coming from horizon scanning of the new technologies expected in the next 5 years.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10	284	In the "Points of attention for the assessment of scoping process" box: Please define "well established" more clearly.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10-11	287-291	A surrogate outcome is a measurement that is not the ultimate/relevant outcome of the disease (biomarker, measure of a function) that has proven to be correlated with the ultimate endpoint of the disease. It should be stated explicitly that any biomarker or intermediate outcome used as a surrogate outcome must have demonstrated such a correlation – not all biomarkers or intermediate outcomes necessarily do so. Further, the biological or other plausibility of the causal pathway from surrogate marker to ultimate endpoint should be stated; otherwise, a statistically significant association between the surrogate outcome and the patient-centered outcome might only indicate confounding. Surrogates may also be used to address issues of confounding in longer term outcomes. E.g., PFS as treatment effect in oncology trials to avoid issues with OS from crossover/subsequent therapies etc.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and	10-11	304-308	This paragraph may be strengthened and better justified by adding "due to the smaller sample sizes and shorter duration often associated with trials employing surrogate markers" to the first sentence. These factors create some uncertainty about whether adverse event risk has been fully captured, which, along with potential greater	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Outcomes Research			uncertainty about clinical benefit, are the main considerations affecting risk-benefit evaluation. However, long treatment duration is not always relevant. Progression free survival is used as a proxy for overall survival in oncology trials and the requirement of a benign safety profile is less applicable to oncology treatments.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10-11	3.3 Surrogate outcomes (Lines 285-336)	Statement in the box, bullet point 1 "If evidence for a patient-centred outcome is likely to be available, then this should be requested during the scoping process instead of surrogate outcomes such as morbidity, overall mortality and HRQoL" is inconsistent with ll.224-225 which names mortality as a patient-centered outcome, as well as with the usual treatment of morbidity, mortality and HRQoL as ultimate outcomes, depending on the condition.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10	287	The source of evidence may not be a trial; we suggest adding “or the available data” to the end of this sentence.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10	297	Must a morbid event be irreversible in order to classify as a significant clinical outcome (eg, exacerbations in COPD)?	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes	10	299	Small point – “rigorously” probably covers the same territory as “rigorously fully” and sounds less redundant.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Research				
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10	308	It would be good if the text could encourage use of rigorously validated surrogate outcomes whenever possible for reasons of speed, ethics and costs	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10	308	In the "Points of attention for the assessment of scoping process" box, please consider adding: The scoping process should consider evidence on the meaningfulness of surrogate measures and how they fit within patient treatment expectations based on patients' direct input.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	10	308-309	Points of attention for the assessment scoping process. Often so-called surrogate endpoints are not intended to replace a patient-centered outcome but are important to supplement or complement the assessment of the patient-centered outcome. They could add value as they could limit gaps related to the certainty of the outcome. For example, PFS, PFS2, time to next subsequent anticancer therapy could be intermediate endpoints that could be informative about the life-expectancy of the patients and are important for treatment decisions. Thus, please consider adding the following bullet: • Surrogate endpoints could be used to complement the added benefit assessment, in particular when such endpoints are labelled in the prescribing information	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and	11	313	Please clarify whether both individual patient-level data (IPD) from clinical trials and meta-analysis of trials are necessary for Level 1 evidence – IPD are not likely to co-reside with evidence from trial level meta-analyses.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Outcomes Research				
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	11	317	More detail on what constitutes 'consistent' in Level 2 would be helpful. Are multiple observational studies (Level 2) necessarily required in all cases?	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	11	323-324	Please clarify whether this means that the HTD should demonstrate the treatment effect in addition to the strength of the association between the surrogate outcome and the patient-centred outcome.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	11	327-328	Regarding the inclusion of immature data: • Please clarify if the requirement to provide data could potentially interfere with the Data Monitoring Committee responsibilities. • This may not be feasible to provide for all outcomes. • How will immature data be used in the JCA process? • Please provide more specific criteria for “immature data,” for example, minimum follow-up time. Should the data be blinded? • This could lead to presentation of small samples with limited interpretations.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes	11	331	Section “Uncertainty- Requirements of JCA Reporting” Please reiterate that the assessor should also report evidence on patients’ direct perception of meaningfulness of the surrogate outcome. Please clarify if prior evidence of surrogacy in the same MOA is considered valid or applicable.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Research			"An indication of whether or not a patient-centered outcome is likely to be available at a later date.": This is likely a reference to OS. Is it worth assessing available evidence on the surrogate alongside other available patient-centered outcomes such as HRQoL and Tx satisfaction (most likely available, and these will demonstrate patient perception of tolerability and potential symptom/ HRQoL improvement) while evidence on other patient centered outcome such as OS are awaited? "In cases for which the association between the surrogate outcome and the final patient-centred outcome has previously been examined but for a different disease stage, population or intervention, the assessment report should consider the implications for the validity of this association in the current population and intervention of interest.": Please clarify if prior evidence of surrogacy in the same MOA is considered valid or applicable.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	11	334-336	Please expand on the frameworks that are cited for assessing surrogate outcomes. More specific guidance about the situations in which they are likely to be considered appropriate and acceptable for JCA would be most useful in this document.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	12	4.1 Terminology for JCA (Lines 337-345, inclusive of box)	In the spirit of patient centeredness, safety aspect from patients' point of view, patient tolerability, needs to be included. Considering the challenges with causality assessment, efforts need to be made to capture safety events that are less frequently observed including SUSARs using standard terminologies (such as PRO-CTCAE). See more elaboration on this point in the comments below.	
D4.4 Endpoints	ISPOR – The Professional	12	4.1 Terminology for JCA (Lines	The document recommends not to use the term tolerability. However, some symptomatic AE (non-severe	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Society for Health Economics and Outcomes Research		337-345, inclusive of box)	but sometimes very bothersome for the patients) are sometimes named under the term "tolerance profile or tolerability". Moreover, FDA and EMA are asking for a more systematic report by patients of symptomatic AE (using the PRO-CTCAE in oncology) in clinical trials, to complement the report by physicians (Safety) but also to describe the tolerability of the new therapy by patients. Repeated studies have shown that for symptomatic AE, there is a disagreement between patients and physicians not in the presence/absence of AE, but in terms of severity and impact on daily life (i.e. physicians tend to systematically underreport severity and impact of symptomatic AE, e.g. nausea compared to patients)	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	12	4.1 Terminology for JCA (Lines 337-345, inclusive of box)	It may not be necessary to mandate to use one term instead of another. Some of these terms do mean different things and cannot be used interchangeably to many people, such as “adverse event”/ “adverse reaction” vs. “side effect”, and “safety” vs. “tolerability” (search on web and you will see their differences). If the real purpose here is to mandate certain variables to be reported instead of putting restrictions on terminology, providing clear definitions of adverse event and safety might serve better.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	12	4.1 Terminology for JCA (Lines 337-345, inclusive of box)	While we agree with avoiding the use of diverse terminologies for the same concept, sometimes terms do differ. Consider that “adverse events are unintended pharmacologic effects that occur when a medication is administered correctly, while a side effect is a secondary unwanted effect that occurs due to drug therapy”.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and	12	342-344 and 371-372	Adverse reactions etc. are commonly part of the Common Technical Documents (CTDs). Avoiding the use of this terminology will be inconsistent with the CTDs. Later in 371-372 it says there are exceptions. More clarification around the use of these safety terms is needed.	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Outcomes Research				
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	12	4.2 Safety: overall and specific adverse events (Lines 346-354, inclusive of box)	Along with MS defining their required safety outcomes, both general and specific, it would be helpful if levels of severity of interest for these outcomes (in addition to the normal definition of “serious AE”) are clearly defined as well.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	12	4.2 Safety: overall and specific adverse events (Lines 346-354, inclusive of box)	Do overall safety results (i.e., all AEs combined) mean % patients with one or more AEs here? If yes, it may still be useful to report incidence of each AE. Assessors may need to see the list to decide which AEs are more important, especially for innovative health technologies.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	12	354-355	In the “Points of attention for the assessment of scoping process “box: In the spirit of patient-centeredness, one shouldn't discount the patient voice in treatment safety. As such, the term 'patient tolerability', used in this context, should not be excluded.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	12-13	4.3 Information to be reported for safety outcomes (Lines 355-380, inclusive of box)	It may be useful for the document to introduce the notion of self-reported AE (using e.g. the PRO-CTCAE).	
D4.4 Endpoints	ISPOR – The Professional Society for	12-13	4.3 Information to be reported for safety outcomes	There is a risk of bias in the causality determination in blinded studies, because blinding is rarely perfect (e.g., difference is AE or clinical response may give the	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Health Economics and Outcomes Research		(Lines 355-380, inclusive of box)	investigator an idea of whether a patient is on control or experimental treatment); the bias is just greater in unblinded studies. Since the conclusion applies to both blinded and unblinded studies, the second sentence does not help. Consider the following alternative statement: "Causality (attributability) between a heath technology and an AE could be described by many terms and scales. However, there is always uncertainty and risk of bias in of the determination of "causality", so all safety outcomes must always be reported, irrespective of causality designation."	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	12	368-369	The guideline requires SUSARs (Any suspected unexpected serious adverse reaction) to be reported. A rationale is not given if/how SUSAR reporting could inform relative effectiveness assessments. In principle, SUSARs address pharmacovigilance questions and are a reporting obligation to health authorities by HTD. SUSARs will undergo rigid evaluation by HTD to decide whether a specific event is a new safety signal. A specific event that had been reported as SUSAR in the beginning of a study need not necessarily to be reported at the end of trial following the safety evaluation. In summary the SUSAR evaluation are reflected in the current drug label. Serious adverse events are reported by the HTD in the HTA dossiers anyway. Please consider deleting the following sentence: Any suspected unexpected serious adverse reaction (SUSAR) should be reported, even if these are (by definition) not requested during the assessment scoping stage.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	13	373-374	Please clarify if these are 'drug discontinuations', 'trial discontinuations' or both	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	13-14	5.1 Definitions and general considerations (Lines 381-405, inclusive of box)	It may be helpful to rephrase the statement regarding PROs being "less objective" to avoid suggesting that PROs are less valuable compared to technological or performance measures. Some would say the major value of including PRO's is to capture the perspective of the patient in a subjective sense. In addition, ObsROs and ProxROs may be worth mentioning in this section, as in some cases patients may be unable to complete PROs.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	13-14	5.1 Definitions and general considerations (Lines 381-405, inclusive of box)	1st bullet – table at top of page 13: While "medical technology" can be considered to cover passive measures such as digital health technologies, please consider including passive technologies in any more detailed discussion about medical technology as a data source.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	13-14	5.1 Definitions and general considerations (Lines 381-405, inclusive of box)	Table at top of p.13 - in addition to the "main source of information" (prior validation evidence package), there should be a conceptual framework to show why the concepts (symptoms/ impact/HRQoL) are identified to be covered by the PROMs selected in the trial.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	13	399-400	It should be clear that the patient is not always able to directly provide their perspective (e.g., young children, patients with some cognitive or physical disabilities). In these circumstances, a nonprofessional observer can complete a questionnaire based upon observed manifestations of specific symptoms or impacts (ObsRO) or a professional assessment based upon specific tests that also require an element of judgment to arrive at the score (ClinRO).	
D4.4 Endpoints	ISPOR – The Professional Society for Health	14-15	5.2 Validity and reliability of scales (Lines 406-458, inclusive of box)	It would be beneficial to mention other considerations for PROMs, including fitness for context of use and the potential relevance of conceptual frameworks to guide the selection of appropriate concepts. In addition, the use of	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Economics and Outcomes Research			the term 'content validity' could be helpful as well, given its importance in overall PROM validity. There are several well-known and well-cited books that may be useful to cite to guide readers to helpful references, specifically Cappelleri et al (2016) and Streiner et al (2015).	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	14-15	5.2 Validity and reliability of scales (Lines 406-458, inclusive of box)	7th bullet – table: For PROMS this description should also include evidence that the instrument is fit for the context of use.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	14-15	5.2 Validity and reliability of scales (Lines 406-458, inclusive of box)	Table on p. 15, last bullet: Regarding the following sentence: "References, as provided by the HTD, allowing the access to the specific (clinical) studies assessing the measurement properties (and measurement model) of the instruments that are used." Please clarify what kind of references should be provided (published, data on file).	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	14-15	5.2 Validity and reliability of scales (Lines 406-458, inclusive of box)	"Content validity" is not discussed. Psychometrics (eg, l. 432) is important to ensure the statistical properties of a questionnaire, but the first step is to be sure 1/ that the content of the questionnaire is really measuring the concepts it intends to capture and 2/ that the concepts are relevant for the patients included in the study	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	14-15	5.2 Validity and reliability of scales (Lines 406-458, inclusive of box)	There are a few other noteworthy, well-cited books that deserve to be referenced: Cappelleri JC, Zou KH, Bushmakin AG, Alvir JMJ, Alemayehu D, Symonds T. 2013. Patient-Reported Outcomes: Measurement, Implementation and Interpretation. Boca Raton, Florida: Chapman &Hall/CRC Press. Fayers PM, Machin D. 2016. Quality of Life: The Assessment, Analysis and Reporting of Patient-Reported	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
				Outcomes. Third edition. Chichester, United Kingdom: John Wiley & Sons Ltd. Streiner DL, Norman GR, Cairney J. 2015. Health Measurement Scales: A Practical Guide to Their Development and Use. Fifth edition. New York, NY: Oxford University Press. Line 422: There can be systematic error as well as random error, so I might drop (i.e., random error) at the end of the sentence.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	14-15	5.2 Validity and reliability of scales (Lines 406-458, inclusive of box)	In empirical studies, the reliability of the measure can place limits on the empirical validity results, meaning that in some cases establishment of reliability may need to precede some aspects of the validity work.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	15-17	5.3 Interpretability of scales (Lines 459-521, inclusive of box)	It would be helpful to clarify the current language regarding MID/MCID to more clearly delineate between group-level and patient-level. In addition, the recent FDA PFDD Guidance 3 workshop introduced the term “meaningful within-patient change (MWPC),” which may be relevant to note. There are also references on responder analyses minimal differences that may be worth citing, such as the recent FDA guidances, Revicki et al (2008), King (2011) and Coon & Cappelleri (2016).	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	15-17	5.3 Interpretability of scales (Lines 459-521, inclusive of box)	There is a recent paper in Value in Health which is useful for this section. To help interpret composite endpoints which use a responder definition (Lines 447-484), it shows how you can link the response criteria to changes in health utility values used by health technology assessment (HTA) bodies. Producing this evidence will help HTA agencies interpret whether the endpoint corresponds with health gain valued according to their preferred instrument (which varies between member state jurisdictions). Paper details here: https://doi.org/10.1016/j.jval.2022.07.001	
D4.4	ISPOR – The	15-17	5.3 Interpretability	Several important references on interpretation are	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
Endpoints	Professional Society for Health Economics and Outcomes Research		of scales (Lines 459-521, inclusive of box)	missing and merit citation: Cappelleri JC, Bushmakin AG. 2014. Interpretation of patient-reported outcomes. Statistical Methods in Medical Research 23:460-483. Coon CD, Cappelleri JC. 2016. Interpreting change in scores on patient-reported outcome instruments. Therapeutic Innovation & Regulatory Science 50:22-29. Coon CD, Cook KF. 2018. Moving from clinical significance to real-world meaning: methods for interpreting change in clinical outcome assessment scores. Quality of Life Research 27:33-40. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. 2007. Understanding the minimum clinically important difference: a review of concepts and methods. The Spine Journal7:541-546. King MT. 2011. A point of minimal important difference (MID): a critique of terminology and methods. Expert Review of Pharmacoeconomics and Outcomes Research11:171-184. McLeod LD, Coon CD, Martin SA, Fehnel SE, Hays RD. 2011. Interpreting patient-reported outcome results: US FDA guidance and emerging methods. Expert Reviews of Pharmacoeconomics & Outcomes Research 11:163–169. Revicki D, Hays RD, Cella D, Sloan J. 2008. Recommended methods for determining responsiveness and minimally differences for patient-reported outcomes. Journal of Clinical Epidemiology 61:102-109. Patient-reported outcome measures (PROMs) of an underlying continuous nature should be primarily analyzed as continuous outcomes to detect treatment effect, and responder analyses should be reserved as secondary analyses for enhancing interpretability and for regulatory purposes of PROMs. Here are supportive references: Collister D, Bangiwala S, Walsh M, Mian R, Lee SF, Furukawa TA, Guyatt G. Patient reported outcome measures in clinical trials should be initially analyzed as	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
				continuous outcomes for statistical significance and responder analyses should be reserved as secondary analyses. J Clin Epidemiol 2021; 134:95-102. Cappelleri JC. Further reduction in statistical power for responder analysis of patient-reported outcomes with measurement error. Journal of Clinical Epidemiology. 2021; 140:200-201.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	15-17	5.3 Interpretability of scales (Lines 459-521, inclusive of box)	Patient-reported outcome measures (PROMs) of an underlying continuous nature should be primarily analyzed as continuous outcomes to detect treatment effect, and responder analyses should be reserved as secondary analyses for enhancing interpretability and for regulatory purposes of PROMs. Supportive References: Collister D, Bangiwala S, Walsh M, Mian R, Lee SF, Furukawa TA, Guyatt G. Patient reported outcome measures in clinical trials should be initially analyzed as continuous outcomes for statistical significance and responder analyses should be reserved as secondary analyses. J Clin Epidemiol 2021; 134:95-102. Cappelleri JC. Further reduction in statistical power for responder analysis of patient-reported outcomes with measurement error. Journal of Clinical Epidemiology. 2021; 140:200-201.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	15-17	5.3 Interpretability of scales (Lines 459-521, inclusive of box)	Suggest adding the term “Meaningful Within-Patient Change (MWPC)” in the guideline. MWPC has been illustrated in the FDA Guidance PFDD Public Workshop Guidance 3 Discussion Document (fda.gov). The current statements do not distinguish the patient-level change (e.g., MWPC) and the group- level difference (e.g., mean difference between treatment groups). It is suggested to clarify the terms in patient-level or group-level when “MID” and “MCID” are used. It would better to align with the recommendations in the FDA Guidance PFDD Public Workshop Guidance 3Discussion Document (fda.gov).	
D4.4	ISPOR – The	15-17	5.3 Interpretability	FDA recommends the use of anchor-based methods	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
Endpoints	Professional Society for Health Economics and Outcomes Research		of scales (Lines 459-521, inclusive of box)	supplemented with both empirical cumulative distribution function (eCDF) and probability density function (PDF). Please consider adding reference to the recent set of FDA guidances -- https://www.fda.gov/drugs/development-approval-process-drugs/fda-patient-focused-drug-development-guidance-series-enhancing-incorporation-patients-voice-medical.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	15-17	5.3 Interpretability of scales (Lines 459-521, inclusive of box)	Please comment on whether 2 studies (clinical trials) are required to establish MCID and measure the response analysis of achieving MICD by treatment	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	15-17	5.3 Interpretability of scales (Lines 459-521, inclusive of box)	Overall, a recommendation about adequate determination of responders would be useful. Of note, validated, established response thresholds are useful and valuable to enhance the transferability of risk/benefit assessments based on PRO measures to assess relative effectiveness for health technology assessment and to ensure consistent interpretation of PRO effects. A singular threshold of x%-change of the continuous scale range for all instruments is incongruent with previously defined and scientifically established thresholds and is not well-suited for universal implementation. [Reference: Schlichting et al, Is IQWiG's 15% Threshold Universally Applicable in Assessing the Clinical Relevance of Patient-Reported Outcomes Changes? An ISPOR Special Interest Group Report, Value in Health, Volume 25, Issue 9, 2022, Pages 1463-1468, ISSN 1098-3015, https://doi.org/10.1016/j.jval.2022.07.010. Please add some language as follows. “In general, validated and established response thresholds should be considered when defining relevant individual response	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
				thresholds in terms of MID. Anchor based methods to determine MIDs utilizing patient-reported anchors are preferred. In absence of patient-reported anchors clinician reported anchors should be considered."	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	15-16	471-476	Interpretability of scales may vary in case categorical scales are transformed into continuous scales or vice versa. Whether a continuous or categorical scale is used to determine the endpoint of interest strongly depends on the underlying objective. One might be interested to explore the rate of patients who improved / maintained / worsened their symptoms compared to baseline (patient level objective), or the average change compared to baseline (within group perspective). It is difficult to argue which of the associated analyses complements the other as they address different questions. The risk of data dredging can be avoided by specifying key outcomes of interest in advance of the analysis which is typically done in the analysis plans. It is recommended to align key outcomes with the agencies in joint scientific advice meetings. The inflation of type I error rate might not be an exclusive problem here but also for subgroup analyses. Please consider changing the language as follows: "this expression of treatment effectiveness can enhance interpretability. Analysis on the categorical scale could complement the analysis on the continuous scale and vice versa. In addition, to avoid the risk of data dredging and, one measure of treatment effect should be pre-specified in the protocol and statistical analysis plan as a primary analysis (see the EUnetHTA 21 practical guideline "Applicability of evidence: practical guideline on multiplicity, subgroup, sensitivity and post-hoc analyses").	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and	16	485-490	The terms MID and MCID are not interchangeable although they are frequently confused and used as if they were the same thing. Suggest clarifying which you are referring to. The definition in Trooster (2011) may be useful: " The minimal effect that would be meaningful to patients is the minimally clinically important difference	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Outcomes Research			(MCID), the minimal difference that reflects a true improvement (or deterioration) in an outcome is the minimally important difference."	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	16	492-498	In the absence of patient-reported anchors, clinician-reported anchors should be used. Please add the following sentence online as follows: When patient-reported anchors are not available clinician reported anchors could be acceptable.	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	16	494	PGIS (impression of severity) may also be used as an anchor	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	16	499-509	We acknowledge that distribution-based methods are informative to statistically characterize MIDs. However, for a relevant patient-centered outcome, responder definition should be primarily considered anchor-based methods ideally utilizing patient-reported anchors. Please consider changing the paragraph as follows: MIDs are also frequently estimated using distribution-based methods [51]. In contrast to anchor-based methods, only the overall variability in scores is used in distribution-based methods. Thus, they are criticized as they do not explicitly refer to the meaning of the change for patients (51). Two approaches are most common. The first is based on estimation of Cohen's d. (delete the rest of the sentence in the guidance document)	
D4.4 Endpoints	ISPOR – The Professional Society for Health	17-21	6 References (Lines 522-674)	The link to the EUnetHTA endpoints guideline (ref no. 14) is not valid. It should be https://www.eunethta.eu/wp-content/uploads/2018/01/Endpoints-used-for-Relative-Effectiveness-Assessment-Health-related-quality-of-life-	

Please add extra rows as needed.

Sub-deliverable	Comment from	Page number	Line/ section number	Comment and suggestion for rewording	Is your comment an editorial comment?
	Economics and Outcomes Research			and-utility-measures_Amended-JA1-Guideline_Final-Nov-2015.pdf	
D4.4 Endpoints	ISPOR – The Professional Society for Health Economics and Outcomes Research	21	677-682	While OS, PFS, EFS etc. are certainly relevant to patients, it should be recognized that these are not specifically considered to be patient centered. As per Reeve at al. (2013), patient centered research is "the integration of patients' perspectives about their health with clinical and biological data to evaluate the safety and effectiveness of interventions. Such integration recognizes that health-related quality of life (HRQoL) and how it is affected by disease and treatment complements traditional clinical endpoints such as survival or tumor affected by disease and treatment complements traditional clinical endpoints such as survival or tumor response in cancer."	

Please add extra rows as needed.