EUnetHTA JA3 WP4 - Other technologies, OTCA19
Review by external experts of the 2nd draft Project Plan for "Screening for osteoporosis in the general population"

November 2018

EEXTERNAL EXPERTS

Comments were received from:

Name	Affiliation
Rafael Azagra	Universitat Autònoma de Barcelona
Johannes Flechtenmacher	Ortho-Zentrum, Karlsruhe
Alexander Mann	Endokrinologikum, Fankfurt am Main
Marta Zwart Salmerón	Institut Català de la Salut, Girona

Comment from *Insert your name and organisation*	**Page number** *Insert 'general' if your comment relates to the whole documen t*	**Line/ section number**	**Comment and suggestion for rewording** *Please insert each new comment in a new row.*	**Character of comment** ● 'major'[a] =1 ● 'minor'[b] = 2 ● 'linguistic'[c] =3 *Please indicate your choice by writing the according number in this field, e.g. for major choose "1".*	**Author's reply**
Zwart	4	List	Rewrite the second column uniformly (lowercase and uppercase). Add other abbreviations that appear in the draft: - CE	3	Reviewed We have added CE = Conformité Européenne (i.e. European Conformity) to the list of abbreviations.
Azagra-Zwart	6	Contributors	Rafael Azagra is Specialist in (1) Family and Community Medicine and (2) Clinical Pharmacology Marta Zwart is Specialist in Family and Community Medicine	2	amended
Zwart-Azagra	6	7, Table 1-2	We find insufficient to consider answers from 3 German patients based on the HTAi questionnaire from patient's organizations. We consider that patients' organizations representatives in the osteoporosis field are very relevant but should not be the unique	1	Thank you very much for highlighting this important point again. Each participating patient filled out a DOICU form and no conflict of interest was identified. Furthermore, when selecting the patient organisations, we made sure that financial

Please add extra rows as needed.

[a] "major": the comment points to a highly relevant aspect and a thorough answer is expected from the author(s)
[b] "minor": the comment does not necessarily have to be answered in a detailed manner
[c] "linguistic": grammar, wording, spelling or comprehensibility

			source and can be very influenced by clinical societies and industry. Even the patients from KOL can be biased because they are patients that already very selected. Also, for example, we don't see any relationship between the feelings of the patients along Europe and having a high prevalence or having European guidelines. So, we think that having patients from more than one country is a very relevant opportunity for an European assessment like this. At least, it could be recommendable to have patients answering the HTAi questionnaire not only from patients' organizations, would provide a more accurate work from a patient engagement perspective.		support from commercial enterprises was made transparent by the organisation (e.g. on their homepages). In the case of the patient organisation whose patients answered our questionnaire, the sponsoring share amounts to about 3% of the total organisational budget. In general, the influence of patients in this assessment is limited to the identification of patient-relevant outcomes. In addition to the consultation of literature, contact with patients should ensure that no patient-relevant outcome is overlooked in the assessment. In our opinion, the effects of the intervention are comparable in all European member states. There may be differences among European healthcare, but it is neither possible nor intended to identify and cover endpoints resulting from different health systems across the board. Instead, it can be assumed that clinical endpoints have very similar importance to all European patients. Of course, it could be assumed that there can generally be a wide range of patient preferences - regardless of the country. However, a procedure would then have to be chosen that actually generates a representative sample of patients from all European countries. For time reasons, this does not seem feasible in the context of a rapid REA.
Zwart	8	Project approach	Add "please see Section" before the specified (4.2) in order to understand the reference.	3	This is not a reference, but version 4.2 of the Core Model for Rapid REA. We added "Version" to make this clearer.
Zwart	8	Project approach	Order the sentences first the one that references the appendix 1 and then 2.	3	accepted
Zwart	8	Project approach	HTA Core Model R for rapid REA must be referenced as in the previous sentence.	3	amended

Please add extra rows as needed.

[a] "major": the comment points to a highly relevant aspect and a thorough answer is expected from the author(s)
[b] "minor": the comment does not necessarily have to be answered in a detailed manner
[c] "linguistic": grammar, wording, spelling or comprehensibility

Johannes Flechten-macher	8	29	The study you propose a limited to the evaluation of the FRAX screening tool only. There are other screening tools in Europe for osteoporosis. In Germany FRAX is generally not used but the DVO guidelines, which differ significantly from FRAX. The algorithm of FRAX is not published. The use of FRAX in a clinical setting is depending on the availability of the program which produces costs. No general answer to the question if screening for osteoporosis makes sense will be possible.	1	Thank you very much for your valuable comment on FRAX. In fact, we have only listed FRAX as an example of a clinical risk assessment tool. It is not intended to evaluate FRAX as a brand, but rather to investigate whether the use of risk assessment tools can be of benefit to the general public. This means that other risk assessment tools (like "any paper-based or electronic approach/instrument that compiles/consolidates various demographic or clinical characteristics of an individual and compares an individual's characteristics against a threshold or guideline to make a subsequent decision for testing or treatment", [AHRQ Publication No. 15-05226-EF-1 June 2018]) are also considered if they have been investigated in appropriate studies. In addition, there are freely available FRAX versions in paper form and online (even in Germany), so we feel that the commercial interest is only moderately present here.
Mann	8		As risk assessment tool only the Frax-score is mentioned (e.g.). What other scores will be evaluated ? It is suggested to also mention the alternative DVO-Score, which is more widely used in German speaking countries.	2	
Zwart-Azagra	9	Clinical effectiveness: - RCT	The aim of this review is to summarize results and assess the risk of bias in included papers and heterogeneity. As the stakeholders explain, a subgroup analysis and investigation of heterogeneity is going to be done. In line with this statement I think is important to include RCT and observational studies too and when evaluating the effects stratify them by study design (RevMan software allows it). Therefore would not be a problem if previously this level of evidence has been decided in other population screenings (mammography, lung cancer,...) because you would have this evidence alone afterwards and would decide what is better to show in this specific health problem.	1	The decision not to include non-randomised studies (NRS) has been critically commented on by several reviewers. Therefore, we present our arguments in detail: 1. Scientific validity: According to the EUnetHTA guideline on NRS, three possible reasons can lead to the inclusion of NRS: a) Organisational reasons preventing RCTs, b) very large effects, or c) external needs to offer a 'best guess' answer. Screening for osteoporosis can be studied a) in RCTs without organisational problems, b) it has at best only a modest effect size, and c) population-based interventions require robust evidence rather than a 'best guess'.
Azagra-Zwart	9	Risk of bias tool	Using The Cochrane risk of bias tool you can assess risk of bias in included studies with respect to random sequence generation and allocation concealment or blinding of outcome assessment in RCT and observational studies.	1	

Please add extra rows as needed.

[a] "major": the comment points to a highly relevant aspect and a thorough answer is expected from the author(s)
[b] "minor": the comment does not necessarily have to be answered in a detailed manner
[c] "linguistic": grammar, wording, spelling or comprehensibility

Zwart-Azagra	9	29, Table 2-2	We disagree with the statement on the SCOOP trial. Without under-rating of course the high relevance of this or any trial, its' a mere existence, justifies not to consider other type of evidence. The circumstances, the object assessed in different screening studies can be different, so the clinical characteristics of the patient especially in our current assessment case, where epidemiology and current strategies available to be assessed are different and discussed among the countries. In this circumstances, the rule of only considering trials because –of course- observational studies have less certainty do not fit from our opinion. This is also the reason why GRADE quality assessment is expected, to say that some conclusions are based on some degree of evidence, and some with less, if needed. Therefore, using only trials because of the "mere" existence of a trial seems a tailored analysis to only use a little number of studies. Indeed, we would not be able to use the assessment in Spain properly if only trials are considered. Observational quality studies on DXA, qCT or qUS, FRAX and Q-Fracture must be taken in consideration. In conclusion, if the assessment is not opened to no-type of observational study, we recommend learning from the disagreement experience, and to use this assessment as a consensus reference, including the following statement below the Table 2-7 as a comment for the study design description "All authors agree that clinical trials are a paradigm to analyze evidence. Nevertheless, Member States could occasionally consider an additional relevant observational study as of an addition-value for national decision-making in case of a potential adaptation"	1	2. <u>Unpredictability of NRS results</u>: While it is possible to assess the risk of bias, the overall direction of the risk of bias for the effect estimates is not known and can go in any direction thereby severely limiting the usefulness of NRS. For screening studies in particular lead time bias and length time bias can only be avoided by using an RCT design. From a methodological point of view it is inappropriate to pool RCTs and NRS because the results presented differ: RCTs report effect estimates and NRS adjusted effect estimates (i.e. one tries to adjust for confounding). In view of this a statistical assessment of heterogeneity is not indicated either. Even the similarity of RCT and NRS results in a split meta-analysis does not prove the correctness of NRS results. 3. <u>Transferability of NRS results</u>: Theoretically, NRS results might be used to extrapolate the RCT results obtained in women to men. However, due to the differences in disease prevalence and life expectancy, such extrapolations are scientifically insufficient. Furthermore, NRS results do not generally have higher external validity (see EUnetHTA guideline). Therefore, NRS findings should not be euphemized as "real world evidence". 4. <u>Lack of consequences</u>: Additional results stemming from NRS would only allow a very low certainty of results (at best "low" according to GRADE). In our view, it would be potentially harmful if Member States use such uncertain conclusions for national decision-making. If RCT data are lacking for screening in men or for certain types of screening interventions, such "white spots" on the map can be described
Azagra-Zwart	10	Subgroup analyses and other effect modifiers: SEX	Since the most of studies recruit postmenopausal females, again all RCT, cohort studies and case-control studies are needed and afterwards analyze and discuss this common methodological heterogeneity of reviews.	1	
Johannes Flechtenmacher	9		You propose that only marker-based strategy design RCTs should be evaluated. To my knowledge just two recent studies fit		

Please add extra rows as needed.

[a] "major": the comment points to a highly relevant aspect and a thorough answer is expected from the author(s)
[b] "minor": the comment does not necessarily have to be answered in a detailed manner
[c] "linguistic": grammar, wording, spelling or comprehensibility

			these criteria. In these studies only FRAX was evaluated. It might be wise to reconsider the study plan,		but should not be masked by filling them with NRS data. Previous examples show that – in the absence of RCT evidence – screening programs in adults will not be implemented, for good reasons. 5. Workload: Searching for, assessing and extracting NRS would require enormous resources and would thus delay the project. Given the uncertainly of NRS results, this workload appears unjustified. 6. No comparison between screening tests: The project aims to answer the question whether screening is beneficial. If one screening intervention is found to be beneficial, we will not examine, whether any of the screening tests in the beneficial screening intervention could be replaced by another screening test. For example, if FRAX is found to be part of an effective screening strategy, we will not check whether any other diagnostic instrument offers similar test accuracy as FRAX. The same applies to heel ultrasound as a possible replacement for DXA. Such questions would have to be addressed in a second assessment. In principle, a EUnetHTA assessment currently has no binding effect for a national uptake. Member States are free to expand the evidence assessed in the EUnetHTA assessment. From our point of view, however, this does not seem to be a discussion that should be addressed within the framework of the project plan, but only takes effect within the framework of the preparation of a reimbursement guideline. We therefore consider it inappropriate to add a
Mann	9	Sec.1, par 1	It is not clear, what type of studies will be included – only marker-based RCTs ? How many studies of that type exist ?	1	
Mann	9	Sec.1, par 1	Why are real world evidence studies excluded a priori ? Please discuss.	1	

Please add extra rows as needed.

[a] "major": the comment points to a highly relevant aspect and a thorough answer is expected from the author(s)
[b] "minor": the comment does not necessarily have to be answered in a detailed manner
[c] "linguistic": grammar, wording, spelling or comprehensibility

					passage such as the proposed one in the project plan.
Mann	9	Sec.2	"the current german guidelines " - this may be misunderstood: the current DVO consists of the german speaking societies in Germany, Austria and Switzerland, please clarify	2	Thank you for this clarification. To our regret, we have called the quoted (among others) guideline "German Guideline", although it is the guideline of German-speaking osteologic societies. It therefore represents the guideline of several European countries and not only of the country of the first authors. Nevertheless, we have completely deleted the quotation at this point due to political reasons.
Mann	16	Sec. 2	Replace german guideline by DVO-guideline	2	
Zwart-Azagra	9	29, Table 2-2	We consider that German guidelines can be of a high interest, but should not be cited differentially (unless selected by given quality criteria), even if authors are from Germany. The assessments should be done with the sight of an European perspective, and therefore, if national guidelines are to be considered, at least all authors, co-authors and reviewers guidelines should be considered, or at least more than one country guidelines should be considered.	1	
Zwart	11	EFF and SAF domains	- I do not understand the cut-point from 2013.	1	The rationale for choosing this cut-off is the following: We aim to identify SR / HTA not older than 5 years. From our point of view this is a reasonable time span for the identifcation of SR / HTA.
Zwart	12	Sources of information retrieval...	- It is necessary the review of the search strategy for identification of studies, both PubMed and EMBASE.	1	Thank you very much for this note: The search strategies will be peer reviewed with PRESS checklist by a second information specialist (from IQWiG).
Zwart	13	I1	- Adults, age 18 years or older.	2	In the end, it will make no difference, whether the literature search is restricted to adults, as osteoporosis is a disease of the elderly. For brevity's sake, we chose not to discuss children here. Indeed, we identified a fair number of studies on children with osteoporosis but these tend to be secondary to other diseases, which would not correspond to the "general population" to be examined in this assessment.

Please add extra rows as needed.

[a] "major": the comment points to a highly relevant aspect and a thorough answer is expected from the author(s)
[b] "minor": the comment does not necessarily have to be answered in a detailed manner
[c] "linguistic": grammar, wording, spelling or comprehensibility

Zwart	13	I6	- Languages?	1	We have deliberately not made any language restrictions, even if it can be assumed that large and high-quality studies will primarily be published in English. However, if there is a suitable study published in a language other than English, it should be taken into account and its transferability checked and discussed if deemed relevant.
Zwart-Azagra	13	Table 2-4	I am worried about the item I2, and the recurrent use of "osteoporosis". Even if the concept is used in the title of many documents related with this assessment, I foresee "fragility" and "fracture" as the concepts used by experts nowadays. Therefore, even if in the Table 2-7 those concepts are clearly included, I would like to see the search engine specific "search strategy" when this are designed, and recommend to be careful to not lose studies that do not use the word "osteoporosis".	1	The search strategy will be developed on the base of a testset. There will be a peer review with PRESS checklist by a second information specialist. If you have relevant references to test the search strategy with, please let us know and send us the PMIDs of these references for a check.
Mann	14	Sec.1	Some substances are written with a capital letters, others not.	3	amended
Zwart	16	Outcomes	Some studies will register non symptomatic lumbar spine fractures too.	1	Outcomes to be assessed in this assessment must be patient-relevant. Explicitly asymptomatic fractures are by definition not patient-relevant and will therefore not be considered.
Zwart-Azagra	20	5.1.	'Model for Rapid Relative Effectiveness Assessment'. Additionally, assessment elements from other HTA Core ModelApplications must be referenced as in Project approach. HTA Core ModelApplications better as 'HTA Core ModelApplications' taking into account the previous style.	3	Thank you for that hint. Unfortunately, this sentence is part of the template and the suffix "if deemed relevant" should surely express that assessment elements from other HAT Core Model Applications are only used when needed. Since this formulation is part of the template, no changes are made.
Zwart	20	A0020	OR lowercase	3	Thank you, we will forward this to the template managers.
Zwart-Azagra	21-22	A002-A005, A0024-A0025, D0001-D0005, D0006, D00011-D0012, C0008	Osteoporosis/fragility fractures in every line.	1	Thank you very much for this valuable comment. For space reasons, we did not follow the suggestion to supplement "fragility fracture" as a supplement to osteoporosis in every relevant assessment element.

Please add extra rows as needed.

[a] "major": the comment points to a highly relevant aspect and a thorough answer is expected from the author(s)
[b] "minor": the comment does not necessarily have to be answered in a detailed manner
[c] "linguistic": grammar, wording, spelling or comprehensibility

					However, we have added a paragraph in the "Approach and Method" section explaining the use of the terms.

Please add extra rows as needed.

[a] "major": the comment points to a highly relevant aspect and a thorough answer is expected from the author(s)
[b] "minor": the comment does not necessarily have to be answered in a detailed manner
[c] "linguistic": grammar, wording, spelling or comprehensibility