NHS R&D National Coordinating Centre for Health Technology Assessment (NCCHTA)
www.hta.ac.uk
nhsnet: nww.hta.nhsweb.nhs.uk

at the
Wessex Institute for Health Research and Development
Mailpoint 728, Boldrewood
University of Southampton
Bassett Crescent East
SOUTHAMPTON SO16 7PX

Applicability Testing of WP5 toolkit - round one response summary report, June 2007, updated December 2007

Table of Contents

Abstract

This report details the responses to round one applicability testing of the WP5 Adaptation Toolkit (Version 2).

The toolkit was developed based on the results of a questionnaire, in-house thinking on content and function, and a Delphi survey of 19 HTA agencies.

Evaluators from 16 European HTA agencies selected one or more HTA reports from a different country and tested the WP5 Adaptation Toolkit as an aid to adapting the report to meet the needs of their own health service. They completed a specially designed qualitative evaluation sheet. Responses were submitted in May / June 2007 (and one in November 2007).

Three of the evaluators took part in 1 hour face to face or telephone interviews to further explore their comments on the evaluation sheet.

Introduction

The WP5 Adaptation Toolkit is one of the deliverables of WP5, designed to help HTA organisations/networks adapt HTA reports from another country for their own use. The toolkit has two objectives, to:

(1) Enable the critical appraisal of reports,
(2) Provide advice to aid adaptation.

Version 2 of the toolkit was applicability tested by European HTA agencies. 27 agencies were approached. 16 Agencies completed a semi-structured evaluation sheet for 17 HTA reports (one agency used the toolkit on 2 reports). This outlined their views about the experience of applying the toolkit in the adaptation of an HTA report from another country for their local needs.

Three of the evaluators took part in 1 hour face to face or telephone interviews to further explore their comments on the evaluation sheet.

Aims

The aim of round 1 of the applicability testing was to allow HTA organisations to try out Version 2 of the Adaptation toolkit, by using it to adapt a single HTA report produced in another country to their own setting, and evaluating the toolkit for this purpose.

Method

Participants

A total of 27 European HTA agencies were contacted, of which 24 agreed to help. 17 evaluations were submitted by mid-June 2007 from 16 agencies. 4 interviews were carried out between June and August 2007.

Survey Questionnaire

The Applicability testing round 1 questionnaire consisted of ten questions:

1. How long did it take you to use the toolkit?
2. Did you use the speedy sifting section to assess the relevance of this report to your question? If so, how useful was it?
3. How can we improve on the speedy sifting section? What additional questions or resources would help you assess relevance?

4. Which domains in the main part of the toolkit did you use for this report? (Technology use & development; Safety; Effectiveness; Cost-effectiveness; Organisational aspects)
5. Can we improve on the checklists within these domains? Is the balance of questions right? (too superficial/too in-depth)
6. What additional toolkit questions and resources would help in adaptation?
7. Did you use the glossary? If so, was it useful?
8. Did you consult anything other than the toolkit e.g. resources, checklists to help you adapt this report?
9. What additional work was required to adapt this report for your target setting (your local context)?
10. Is there any other information you would like to provide us with to help improve our toolkit?

See Appendix 1 for the full length version of the survey questionnaire.

Interviews

1 hour interviews carried out face to face or by telephone were carried out with questionnaire respondents in order to build on responses from applicability testing evaluations.

See Appendix 3 for the interview schedule.

Process

The draft toolkit was developed in April 2006. In May 2006 the draft was circulated to all 29 WP5 partners with round 1 of the Delphi survey questionnaire. The WP5 partners' responses were fed into the toolkit.

In March 2007, 27 WP5 partners were approached and asked to participate in round 1 of the Applicability Testing of version 2 of the toolkit. They were asked to nominate an HTA report from another country that they would like to adapt. They were asked to apply the toolkit and complete a semi-structured evaluation sheet consisting of qualitative questions. They were given 1 month to complete the exercise.

In June 2007, 5 respondents who had completed evaluation sheets for round 1 of the Applicability Testing of version 2 of the toolkit agreed to take part in follow-up interviews. 4 interviews were completed by August 2007.

Data Analysis – Survey Questionnaire

Response rates, technology type and country of origin of reports were summarised quantitatively, using descriptive analysis: percentages, ranges and Likert scale-like interpretation where appropriate (all 10 questions).

Qualitative data were analysed using a thematic approach. Themes were identified by considering all responses, drawing out patterns, similarities and differences, and interpreting meanings across the responses. For questions where there was a lack of convergence of perceptions, this is stated in the text (all questions except Q.10).

The results from each question are detailed below.

Data Analysis – Interviews

Responses were analysed using a thematic approach. Themes were identified by considering all responses, drawing out patterns, similarities and differences, and interpreting meanings across the responses. The number of respondents informing each theme is stated in the text.

Results

Response Rates

Twenty-seven organisations / networks were approached. This included 1 LP, 19 APs and 7 CPs. Twenty-four agreed to participate (89%), of which 17 submitted details of a HTA report they wished to use in the test. Of the 24 participants, 20 were APs (83%) and 4 were CPs (17%). Fifteen of the 24 organisations / networks which agreed to participate responded by mid-June 2007, 1 of which tested the toolkit on 2 separate papers (63% response rate, 56% of those originally approached). An additional organisation (DACEHTA) responded in November 2007 (raising the

response rate to 67%, 59% of those originally approached). Of the 16 organisations / networks which responded, 1 was the LP (6%), and 15 were APs (94%). 1 of the APs (NOKC Norway) submitted responses for 2 different reports.

Technology Type used in Applicability Test
Of the 17 HTA reports used to test the toolkit, 5 were about drugs (29%), 1 screening (6%), 4 therapeutic (24%), 2 devices (12%), 4 surgery (24%) and 1 diagnostic (6%).

Country of Origin and language of reports used in Applicability Test
Of the 17 HTA reports used to test the toolkit, 8 were from the UK (47%), 3 from Canada (18%), 1 from Australia (6%), 2 from Belgium (12%), 1 from France (6%), 1 from Sweden (6%) and 1 could not be ascertained (6%).

9 of the 17 HTA reports used to test the toolkit (53%) were in English, 2 (12%) were in French, 3 in English / French (18%), and 2 (12%) not known.

	Country of report							
Country of Responding Agency		UK	Canada	Belgium	Australia	France	Sweden	Not known
	UK		1					
	Austria			1				
	Denmark****	1					1	
	Estonia	1						
	Finland	1						
	France			1				
	Germany							1
	Italy *	1	1		1			
	Norway **	1	1					
	Slovenia	1						
	Spain ***	2				1		
	Totals	**8**	**3**	**2**	**1**	**1**	**1**	**1**

* 3 Italian Agencies responded,
** NOKC in Norway tested the tool with 2 reports,
*** 3 Spanish Agencies responded
**** 2 Danish Agencies responded

Question 1: How long did it take you to use the toolkit?

Quantitative results
17 evaluators responded to this question. 15 of these responses (88%) included a numerical estimation of time.

The range of responses was from ¾ hour to 5 days.

The responses were: ¾ hour, 1 hour, 1 hour, 1 hour + 10-15 minutes, 1 ½ hours, 2 hours, 3 hours + 15 hours, 4 hours, 5 hours, 1 ½ days, 2 ½ days, 3 days, 3 days + 1 ½ hours, 4 days, 5 days.

Textual estimates included: 'quite a while', and 'no comment'.

There was a wide variation in the length of time taken by respondents to use the toolkit. Some of the reasons for this variation can be induced from the themes below. Further reasons can be induced from responses to other questions in the evaluation form:
1) how experienced the respondent is in terms of familiarity with critically appraising HTA reports, with the particular domains that they are evaluating, with critical appraisal tools in general, or due to involvement in the development of the toolkit – eg. the respondent who took ¾ hour 'took 5 minutes' to use the speedy sifting section, and was clearly very familiar with the concepts included in the toolkit: 'those are all question one (HTA-researchers) should know by heart when looking at reviews';

2) how thoroughly respondents undertook the task – eg. the respondent who took ¾ hour 'did not read every single question', the respondent who took 1 ½ days 'passed most economic evaluation questions', whereas the respondent who took 5 days was 'reading the report, answering the questions of the toolkit and annotating what should be changed';
3) how much time the respondent had available to carry out the task – eg. the respondent who took ¾ hour did not use the glossary because it was 'too time consuming', and answered each question in the evaluation sheet very succinctly;
4) it is possible that the respondents' familiarity with the language of the report to be adapted may be a factor in the length of time taken to carry out the task, but there was insufficient evidence from responses to tell whether this was the case.

Qualitative responses
9 of the 17 evaluators (53%) gave further detail about how their estimated time had been utilised (see Appendix 2 for details).

Themes
The following themes were drawn from the qualitative responses:

- The length of time taken depends on the size, structure and complexity of the report evaluated
- The length of time taken depends on how many people are working on the evaluation, and how much discussion is included
- The length of time taken depends on whether users look at every question or every section of the toolkit
- The length of time taken depends on whether the user is already familiar with the HTA report being evaluated
- The length of time taken depends on how many 'problems' are encountered during the evaluation process
- The length of time taken depends on how many aspects of the report are being adapted, eg. only effectiveness, effectiveness and economic evaluation

Question 2: Did you use the speedy sifting section to assess the relevance of this report to your question? If so, how useful was it?

Quantitative results
17 evaluators responded to this question.

All 17 respondents had used the speedy sifting section (100%).

1 respondent described the speedy sifting section as 'very useful' (6%), 8 as 'useful' (47%), 3 as 'quite useful' (18%), and 1 said that it 'worked reasonably fine' (6%). 1 of those who described it as 'useful' also stated that "it is more useful before choosing a report to adapt". 1 respondent stated that "the questions 6, 7, and 8 were the more usefull. 5 could be included in 8." (6%). 1 respondent stated that questions "1-4: important and useful, 5: not crucial for this speedy sifting process."

5 respondents quantified the time taken to use the speedy sifting section. 2 of these had also included the same estimation in answer to Question 1. It is not clear whether the other 3 had included these in the time estimations given for Question 1, or whether this was additional time not accounted for there.

The time estimations for speedy sifting were: 5 minutes, 10-15 minutes, around 15 minutes, 2 hours (including 1 hour to read the text in order to be able to answer the questions), 3 hours.

Qualitative responses
13 out of the 17 respondents (76%) gave further detail about the usefulness of the speedy sifting section (see Appendix 2 for details).

Themes
The following themes were drawn from the qualitative responses:

- The speedy sifting section is easy and fast to apply
- The questions in the speedy sifting section are common sense and those that should naturally be applied by evaluators

- The questions in the speedy sifting section are relevant and reasonable for first appraisal of a report
- Decisions about choosing a HTA report to adapt will have been made prior to use of the speedy sifting section, making the speedy sifting section a formality
- The speedy sifting section allows identification of relevant issues and potential methodological problems with a HTA report
- Usefulness of the speedy shifting section is related to the quality of reporting and structure of the HTA report

Question 3: How can we improve on the speedy sifting section? What additional questions or resources would help you assess relevance?

Quantitative results

16 out of 17 evaluators responded to this question (94%).

11 respondents made suggestions for improvements (65%), 5 stated that no improvements were needed (29%).

Qualitative responses
(See Appendix 2 for details).

Themes
The following themes were drawn from the qualitative responses:

- An additional sifting question is needed, listing the 5 domains, with a grading scale, to allow users to assess how useful each of the domains is likely to be (5 replies informed this theme).

Suggestions each made by only 1 respondent

- The word 'adequate' should be added to question 3: 'Is there an *adequate* description of the health technology being assessed?' and question 8. 'Have the methods of the assessment been *adequately* described in the HTA report?'
- Question 2 should be split, as it currently contains 2 questions
- An additional question is needed: 'Does the population of the study match your own population?'
- The word 'scope' in question 4 needs to be clarified
- An additional question is needed: 'Who was the external reviewer?'
- A judgment is needed to answer question 1, as the policy / research question is not always clearly stated
- An additional question is needed: Is this HTA report based on a systematic and adequately comprehensive literature review?
- An additional question is needed: Was relevant outcomes on safety and effectiveness assessed?
- Given that inexperienced HTA-doers will use the toolkit the speedy sifting section needs to include more questions to be able to decide internally whether to do the adaptation (a list of 31 questions was suggested: see Appendix 2 for details)

Question 4: Which domains in the main part of the toolkit did you use for this report? (Technology use & development; Safety; Effectiveness; Cost-effectiveness; Organisational aspects)

Quantitative results

17 evaluators responded to this question.

The number of respondents using each domain were:
- Technology use & development: 13 (76%)
- Safety: 12 (71%)
- Effectiveness: 16 (94%)
- Cost-effectiveness: 10 (+1*) (59%-65%)
- Organisational aspects: 5 (+1*) (29%-35%)

* One respondent had put these domains in brackets ( ), but no explanation for this was given

No qualitative responses were sought for this question.

Question 5: Can we improve on the checklists within these domains? Is the balance of questions right? (too superficial/too in-depth)

Quantitative results

17 evaluators responded to this question.

11 of the respondents made suggestions for changes and / or clarification (65%). 2 stated that the balance of the questions was right (12%). 1 stated that the questions on cost-effectiveness were too in-depth compared to the other domains (6%), 1 stated that the toolkit is too detailed for internal HTA use, and external experts will be brought in to carry out the adaptation (6%).

Qualitative responses
(See Appendix 2 for details).

Themes
The following themes were drawn from the qualitative responses. There was a lack of convergence of views in the responses, so the themes relate to individual comments:

Overall

- The toolkit does not have enough questions to assess transferability
- Some of the more technical questions need explanatory notes, links to a glossary
- The term 'reliability' could be replaced with the more familiar 'internal validity'
- All question boxes should have the same structure, so eg. the Effectiveness domain questions to assess reliability should be divided into i) aspects of sources of information on effectiveness data, and ii) quality of effectiveness assessment, as is done in the Safety domain
- The questions on relevance, reliability and transferability in each domain are logical and aid the organisation of the main part of the report
- There is overlap in questions in the safety and effectiveness domains, which might be simplified
- There are questions that cannot be applied to every research question, but they are still necessary in the toolkit
- The main part of the toolkit was too comprehensive, and should focus on internal validity
- A reference to the more in-depth resources should be made at the beginning of each section
- The appropriateness of the checklists depends on the expertise and specialty of those using them
- The toolkit was not suitable for the evaluation of an HTA report on screening – the questions on transferability were not relevant - the questions should be more generic
- The questions relating to assessment and transferability should be more clearly separated
- There are many questions, and many of them are long, making it a time-consuming process – a structured questionnaire may be more suitable
- The checklist and resource lists should be more incorporated, with more direct links between them, possibly in a software kit

Section 5.1 – Technology's use domain

- Question 2 should be moved to domain 5.3
- Question 2 should be presented at the end of the assessment process, rather than the beginning
- Question 5 should be moved to domain 5.5
- Question 6 is difficult to comprehend
- Question 7 could take a long time to answer for some technologies due to the need to gather information

Section 5.2 – Safety domain

- Additional questions are needed on searching: 'Was the search for studies reasonably comprehensive and does it include safety aspects?', and / or: 'Was there a specific search for side effects?'

- Question 2 is covered by the speedy sifting process
- Questions 3, 4 and 5 should be listed under a single aspect: "the aspects that should be assessed concerning the sources of information on safety data are:" (2 respondents made this suggestion)
- Question 9: the term 'validity' should be linked to a glossary and defined as internal and external validity
- Question 10 is too in depth – it is difficult to find a report that accomplishes this
- Question 14 is only relevant for drugs

Section 5.3 – Effectiveness domain

- Sub-headings could be added to the reliability section, and relevant questions grouped under them, eg. Searching (Q.4, Q13), Selection of studies (5,6,7)
- Question 1: the first sentence should be removed, and the remainder amended to: 'Are the research questions considered within the effectiveness and efficacy section of the HTA report relevant to your HTA question?'
- Question 1: is covered by the speedy sifting process
- Question 3: should be moved to follow question 4 in the reliability section
- Questions 3 and 4 could be merged
- Question 7: is covered by the speedy sifting process
- Question 9 is difficult to comprehend
- Question 9: does 'all studies referred to in the text' mean included studies?

Section 5.4 – Economic evaluation domain / Cost-effectiveness domain

- This domain is too large
- Some reports provide the results of a literature review, others carry out an original economic analysis: there should be a separate list of questions for each of these 2 scenarios
- Questions 5 and 7 also occur in the effectiveness domain, so could be skipped if the whole report is being adapted
- Questions 11 and 12 could be combined
- Question 23 is difficult to comprehend
- Question 24 is difficult to comprehend
- Question 25 is difficult to comprehend
- Question 27: should be amended to say: 'How generalisable and relevant are the results, the validity of the data and model, to the relevant target setting?'
- Question 28: these elements should be moved to the organisational aspects domain: technological context; personnel characteristics; epidemiological context (including genetic variants); factors which influence incidence and prevalence; demographic context; life expectancy; reproduction; pre- and post-intervention care; integration of technology in health care system; incentives

Section 5.5 – Organisational aspects domain

- Figure 2: (/) is needed between Organizational aspects / dimensions
- Figure 2: 'Inter-organizational level' and 'Intra-organizational level' are difficult to comprehend, and footnotes could be used
- Figure 2: it is unclear what type of organisation is envisaged
- Figure 2: it is unclear what is meant by 'utilisation'
- Additional question needed: reimbursement or no reimbursement

Question 6: What additional toolkit questions and resources would help in adaptation?

Quantitative results

13 out of 17 evaluators responded to this question (76%).

1 respondent noted "no comments". 8 out of 17 respondents did not suggest additional questions and resources (47%).

Qualitative responses
(See Appendix 2 for details).

Themes
The following themes were drawn from the qualitative responses. There was a lack of convergence of opinion in the responses, so the themes relate to individual comments:

- The toolkit should be simplified rather than made more comprehensive
- Standard result extraction sheets, including types of result to be extracted, eg. RR, NNT, ICER, would be useful
- Legal aspects, in relation to how to deal with new technologies would be useful
- The question 'What are the proposals for future research?' would be useful
- Question 14 of the effectiveness domain: 'baseline risk of patients' needs explanation, and the question could be broadened to include the characteristics of the patients and the setting in which they live (socio-demographic and prognostic factors). The word 'risk' could be replaced by 'health status'
- It would be useful to have further questions about the representativeness and generalisability of the trials included in the HTA report. For example, there could be further questions about the patients (e.g. whether the trials were highly selective, or whether they included a broad range of patients and sub-groups); the intervention characteristics (e.g. whether it would be feasible, and acceptable in the target population - particularly important for public health interventions); the setting and organisation (e.g. whether there is adequate capacity and structure to provide such an intervention in the target location). A useful resource for this is: Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity http://www.biomedcentral.com/1471-2288/6/56/abstract
- Question 6 of the economic evaluation domain: 'evidence of the product's efficacy…' the wording should be changed to reflect any technology, not just 'products'
- Resources for the effectiveness domain: this resource is from the early 1990's and may be out of date

Question 7: Did you use the glossary? If so, was it useful?

Quantitative results

16 out of 17 evaluators responded to this question (94%).

6 respondents stated that they did use the glossary (35%), 9 did not (53%), although 4 of these stated that they had read it or browsed it, 1 in their role as a WP5 partner as part of the development of the toolkit. 1 organisation / network used the glossary for the effectiveness domain, but not for the economic evaluation domain (7%).

Qualitative responses
(See Appendix 2 for details).

Themes
The following themes were drawn from the qualitative responses:

- The glossary is very useful to those who are not familiar with the terminology
- The glossary aids understanding
- Use of the glossary is too time consuming
- Additional economic terms are needed, eg. 'relative cost', 'allowance for uncertainty', stochastic sensitivity analysis'
- It would be preferable to have one definition per term in the glossary
- It would be preferable for each glossary term to have an 'official' definition, followed by an example, followed by comments or other definitions
- The inclusion of definitions from different agencies gives a better overview of the term
- The hyperlinks are very useful and time-saving
- There is a typo in the glossary – 'speedy shifting' instead of 'speedy sifting'
- The glossary in the effectiveness domain is clear and exhaustive but rather lengthy

Question 8: Did you consult anything other than the toolkit e.g. resources, checklists to help you adapt this report?

Quantitative results

17 evaluators responded to this question.

8 out of 17 respondents did not consult anything other than the toolkit (47%). 7 respondents specified other resources they had consulted (41%).

Qualitative responses
(See Appendix 2 for details).

Themes
The following themes were drawn from the qualitative responses:

- The resources and checklists listed in the toolkit are useful, usable and familiar, but could be better categorised
- Links to full papers would be preferable to links to abstracts
- The British National Formulary was used to clarify the current licence indication for a drug
- In the effectiveness domain, the 'Critical appraisal worksheet for therapy' resource was used for determining the meaning of 'appropriate criteria' in the assessment of validity of studies
- A rheumatologist was consulted about local treatment strategies and the relevance of a HTA report in the context of that country
- WP4 EUnetHTA Core Model 'utilisation' and 'work processes' sections were consulted, and some errors were identified (more details to follow)
- Internal checklists for assessing systematic reviews were used
- The purpose and meaning of the box of micro, meso and macro levels was not understood
- At least clinical experts, but perhaps also other experts would need to be consulted (eg. in relation to local organisational issues)

Question 9: What additional work was required to adapt this report for your target setting (your local context)?

Quantitative results

14 out of 17 evaluators responded to this question (82%).

Qualitative responses
(See Appendix 2 for details).

Themes
The following themes were drawn from the qualitative responses:

Additional work was needed on:

- How the information relates to / can be supplemented by data from the local context:
    - Epidemiological data: incidence and prevalence information
    - Assessment of the need for the intervention in the local setting
    - Assessment of local patients' risks
    - Local clinical guidelines, drug licensing, policies and current practice
    - Number of times a surgical intervention has been carried out in the local setting
    - Local use of the technology
    - Gathering local cost data
    - Data relating to the local organisational setting
- Updating the literature search on effectiveness, safety, and cost-effectiveness, extending it to include literature written in the local language
- Analysing other reports, clinical guidelines and some original studies on the same topic
- Comparing treatment strategies in the local context with those presented in the report
- Comparing specialties and functions of practitioners in the local context with those presented in the report
- Re-building economic models using locally acceptable cost and discount rates
- Liaising with report authors to ascertain details
- Working and meeting with colleagues, specialists and stakeholders, presenting findings and iterating

Additional themes:

- There are 2 separate processes: i) deciding whether the report *can* be adapted to the local setting, and ii) writing a new report, or summarising the important issues for policy-makers

Question 10: Is there any other information you would like to provide us with to help improve our toolkit?

Quantitative results

13 out of 17 evaluators responded to this question (76%).

4 out of 13 respondents had no further information to offer (31%).

Qualitative responses
(See Appendix 2 for details).

Themes
The following themes were drawn from the qualitative responses:

- The toolkit is extensive and offers good guidance, especially to novices, but is overly extensive for experienced researchers
- The toolkit is of limited value but helps to verify that initial judgments are based on all the relevant aspects
- The toolkit is not much different to other checklists for reviewing literature
- Where there are several HTA reports on the same topic it would be useful to have a list of criteria on which to base a decision to select one of them for adaptation
- Each toolkit domain should ideally be assessed by an expert in that discipline
- Economic aspects:
    - It would be helpful if the toolkit gave information about how to obtain electronic versions of economic models
    - The economic evaluation domain is valid for cost-analysis studies but an additional section would be needed to evaluate reviews of economic studies
    - It will be difficult to adapt cost information when price structure and consumption structure are not comparable
    - Economic evaluation domain, question 28, parameter l) is difficult to comprehend
- Outputs
    - It would be useful to have a template on which to enter notes and conclusions while considering the toolkit questions, which would become the output evaluation report
    - It would be useful to have an electronic summary table of responses to toolkit questions, to enable comparisons between different HTA reports

Interviews

Responses
(see Appendix 4 for details).

5 interviews were arranged with participants from round 1 Applicability testing of the toolkit, of which 4 interviews took place by August 2007.

Themes
The following themes were drawn from the 4 interviews:

- The final product of the evaluation exercise was used in a real life setting, for politicians, healthcare managers, and a directorate of social affairs (2 respondents)
- At present the toolkit is geared towards treatment reports, some questions cannot be applied to screening and further appropriate questions will be needed (1 respondent)
- The toolkit needs to be translated into languages other than English for use by HTA organisations, clinicians, statisticians, engineers, and stakeholders / clients (2 respondents)

- The language of the original report can be a barrier to adaptation for some personnel in HTA organisations (1 respondent)
- The speedy sifting section is easy to follow, and is useful to aid selection of one report among many, or for deciding whether a single report is fit for purpose, although some judgements about availability, quality and relevance were made before applying speedy sifting
- There is a need for an additional speedy sifting question asking whether all the aspects of the topic that adapters need for their adaptation are addressed in the report being assessed, ie. effectiveness, cost-effectiveness, safety, organisational aspects (2 respondents)
- The toolkit is too comprehensive for experienced users, but it would be difficult to decide which questions to cut, so a full and a short version may be useful (2 respondents)
- It would be useful to have example adaptations in the toolkit (1 respondent)
- The Organisational aspects matrix needs updating to match the WP4 core model, and additional useful questions from the WP4 core model could be incorporated into the toolkit (1 respondent)
- The Safety and Effectiveness domains could be merged (1 respondent)
- The Cost-effectiveness domain is very long and complex, and some of the questions need to be simplified (1 respondent)
- Resources listed in the toolkit need to be categorised and referenced (1 respondent)
- Glossary definitions for each term need to be succinct, easily understood by those not involved in HTA, and agreed by members (1 respondent). The glossary is 'messy' (1 respondent)
- The toolkit would be used again as a checklist when adapting reports for important clients, or incorporated into HTA organisation practice (2 respondents)
- An electronic version of the toolkit is needed, which includes example adaptations (1 respondent)

Conclusions

The results of the survey can be summarised as follows:

- Overall the comments were very positive with respect to the usability, usefulness and content of the toolkit, although one or two respondents felt it was too comprehensive, especially for experienced users
- Many areas where meanings could be clarified were identified
- Several changes were recommended to improve the flow / logical progression of questions
- There was a strong feeling (5 out of 17 responses) that an additional sifting question was needed, listing the 5 domains, with a grading scale to allow users to assess how useful each of the domains is likely to be
- The inclusion of output templates on which to collate results of evaluations was suggested by 2 or 3 respondents
- There was a sense that the toolkit was more appropriate for some report types than others – respondents testing screening HTAs and economic reviews found the toolkit did not include all the relevant questions

The results of the interviews can be summarised as follows:

- Interview responses endorsed the conclusions of the survey
- Further suggestions for improvement were suggested

Additionally:

- The importance of translation of the toolkit into languages other than English was emphasised
- Respondents stated that they would use the toolkit again, and recommend it to others to use, particularly if there was an electronic version

Appendix 1: Evaluation sheet with covering letter

eunethta

Dear WP5 members,

Welcome to the first round of our applicability testing!

This document is your evaluation sheet. You should have received a copy of our toolkit (as a word document) along with this evaluation sheet as e-mail attachments. Your task is to adapt one or more HTA reports using our toolkit and to complete this evaluation sheet.

Please choose which HTA reports you would like to adapt and e-mail Hilary Bunce, hilary.bunce@soton.ac.uk, with your decision by Monday 16th April. Then, complete this evaluation form for each individual HTA report you choose to adapt using our adaptation toolkit. For example, if you adapt information from two different reports dealing with the same topic please complete a form for each report (two forms, one for each report). Please e-mail Hilary Bunce (hilary.bunce@soton.ac.uk) with your completed evaluation sheet/s by Monday 14th May. There is no need to submit your adapted report/s to us.

We look forward to receiving your evaluation sheet/s.

If you have any questions or concerns regarding this task please do not hesitate to contact us.

Kind regards,

Debbie Chase
Hilary Bunce

On behalf of the NCCHTA EUnetHTA team

Section A: Information about you & your agency			
HTA Agency		Country	
Your name		Email	

Section B: Information about the report you wish to adapt	
HTA topic	
Report title	
Authors	
Date of publication	
Publisher	
Language of report	
Country of origin	
URL link to report *(if possible)*	

Please limit your responses to each question to 500 words or less

Section C: Information about your experience of using our adaptation toolkit
1. How long did it take you to use the toolkit?

2. Did you use the speedy sifting section to assess the relevance of this report to your question? If so, how useful was it?

3. How can we improve on the speedy sifting section? What additional questions or resources would help you assess relevance?

4. Which domains in the main part of the toolkit did you use for this report? Please tick the relevant boxes

Technology use & development		Safety		Effectiveness	
Cost-effectiveness		Organisational aspects			

5. Can we improve on the checklists within these domains? Is the balance of questions right? (too superficial/too in-depth)

6. What additional toolkit questions and resources would help in adaptation?

7. Did you use the glossary? If so, was it useful?

8. Did you consult anything other than the toolkit e.g. resources, checklists to help you adapt this report?

9. What additional work was required to adapt this report for your target setting (your local context)?

10. Is there any other information you would like to provide us with to help improve our toolkit?

Appendix 2: Individual qualitative responses to evaluation sheet

Question 1: How long did it take you to use the toolkit?

- 5 days: "reading the report, answering the questions of the toolkit and annotating what should be added or changed."
- 1 ½ days: "Reading quickly the report and answering to the questions, it took 1½ days for one person. We passed most economic evaluation questions because our research question does not include such deep analysis on it."
- It took us quite a while: "We have decided to adapt only the effectiveness part, but there occurred some problems, so that it will take longer for adaptation of this report then expected."
- ¾ hour: "but I did not read every single question."
- "Speedy shifting (sic): one and a half hours. Adaptation toolkit: 3 days."
- "The economic domain part was undertaken in 2 working days. The effectiveness domain part was undertaken in another 2 working days by a different user. Some hours have to be taken into account to discuss and revise the answers and the ressults together."
- 1 Hour: "however, report was assessed retrospectively (i.e.report was known to myself already)."
- 5 hours: "just for using the toolkit (HTA report reading hours not included)."
- "10-15 minutes to answer the speedy sifting section. 1 hour to use the main part of the toolkit. The toolkit was used after a first reading of the whole report was completed."
- "Appox 1 h, but failed to make it trough the main part of the toolkit. Speedy sifting process worked fine."

Question 2: Did you use the speedy sifting section to assess the relevance of this report to your question? If so, how useful was it?

- "We have recognized the relevance of the report quite quickly as the report is a very good one. It would be interesting to know what would have happen with a not so good report."
- "The questions are in many respects commonsense, but they are all relevant and reasonably."
- "It was quite useful though the questions were self-explanatory."
- "The questions 6, 7, and 8 were the more usefull. 5 could be included in 8. Actually, the answer to the question 5 will not change my decision to proceed."
- "Good guidance to look critically at the report."
- "It was quite useful. However, I suspect that users of the toolkit will 'pre-select' HTA reports that appear most relevant to them and which are most reliable, prior to doing any formal speedy sifting. Sifting will therefore be a formality – but a necessary one as it will encourage users to be explicit about the reasons for their choice of HTA report."
- "It is good to go through quickly and to obtain a general idea of the report to adapt."
- "we used the speedy sifting section even though we guess it is more useful before choosing a report to adapt. Anyway it was useful as a very first appraisal of the report since it includes the most relevant questions."
- "These questions should be naturally replied by any researcher who is thinking of adapting any report. However, we found this section quite useful, because it provides a fast overview of the HTA report to be adapted and helps clarify the relevant issues. It's eight questions are easy and fast to reply and after replying this section it can be clearly seen if the report is suitable to be adapted. In our case we found the HTA report suitable to be adapted to Spanish context."
- "it was useful to get an overall impression of the report. Using the speedy sifting tool took around 15 minutes (report very well structurated and brief). Usefulness of the toll probably depends on the quality of reporting in the HTA-report."
- "also if the policy/research question was not clearly stated (the executive summary reported only that the purpose of the work was "to conduct a HTA of MoM hip resurfacing arthroplasty")."
- "we used this section, however it was not determinant for our choice. It was useful anyway to have a first general overview of the HTA including some potential methodological problems and it took us just 10-15 minutes to go through this part."
- "Useful and fast sifting process."

Question 3: How can we improve on the speedy sifting section? What additional questions or resources would help you assess relevance?

- "We think about an introductory question about every section to be adapted, something in this way which sections of an hta report do you need to adapt? (technology use; safety; effectiveness; economic evaluation). Adding also the result of using the section of every subject (pass / not pass). In this way you obtain a summary of the process of adaptation over the whole report."
- "You can improve the speedy sifting by using 'structured questionnaire' where you have to thick the boxes (see the next question below). There could be also some kind of graduation or scale with options for example Totally / Partly / Not at all."
- "The questions seem quite comprehensive and I cannot think of anything obvious that has been omitted. Suggestion to add the word adequate to question 3: Is there an adequate description of the health technology being assessed?, also for question 8. 'Have the methods of the assessment been adequately described in the HTA report?'."
- "In Box2- Speedy sifting questions: Assessment of relevance: Question n.2: there are 2 questions in one box and the first question is not possible to answer with "yes/no". Maybe better to place the 1st question in 2nd place and as additional information, not as a question."
- "Since we are evaluating a report in order to adapt it to a defined country setting, a question related to the characteristic of the population involved in the studies may be relevant, such as: Is the population of the study weel described? Are there defined characteristic of the population that affect results of the study? Is it divided into different target groups? To group all these question in one: Can the population of the study match your country's one? Another question that may be can be included is: "Is the answer to the study given?". The study we were adapting comprehended both safety, clinical and cost effectiveness of anakinra for the treatment of rheumatoid arthritis. Although it is a review it could not assess the safety and if we were mostly interested in it, we could have chosen another report from the beginning."
- "Maybe it would be convenient to detail the first question about relevance of the policy and research questions to our question – there can be only some part or parts of the HTA report relevant to our research and therefore it can be difficult to answer the first question simply yes or no."
- "The question "Is the scope of the assessment specified?" may not be clear enough. Probably an explanation of "scope" in the glossary would help. Other terms have been used previous projects like "Aspects". In the EUnetHTA project we are also using other terminology like "Domain". Thus an explanation of "scope" is needed in which its equivalence to the above mentioned terms is aknowledged."
- "I did not found in report information about 5th question of speedy sifting section " Has the report been externally reviewed?" Is it so, that if we have report done by world-famous investigators, then external review is not compulsory? My proposal is in connection with 5th question: if report was externally reviewed to add question about external reviewer: Who was external reviewer? Results of speedy sifting show that this report is very relevant for adaptation in Estonian context. I made this decision within 2 hours and 30 minutes."
- "The first 2 questions of the Speedy sifting are supposed (see p. 7 of the EunetHTA Adaptation Toolkit version 2) to have a kind of gatekeeper function, but this could result to be misleading if considering question 1. This is due to the lack of a standard index for all HTA reports (which implies the not for granted inclusion of a clearly stated policy/reserch question). Although the HTA's research and policy questions are not clear, one can always find useful information in sections dealing with description of technology, technology use etc. to be incorporated in one's own report. May be a judgement is necessary for the first question too, anyway it could be useful to include among criteria for selection of reports the use of a standardized index."
- "Is this HTA report based on a systematic and adequately comprehensive literature review? Was relevant outcomes on safety and effectiveness assessed?"
- "Taking into account that the people who did the applicability testing are inexperienced HTA-doers and that we are commissioning the adaptation - we suggest to expand the speedy sifting section by including more questions to be able to decide internally whether to do the adaptation (see the file attached in the mail)."
- *See also* list of suggested Speedy sifting questions at the end of Appendix 2

Question 4: Which domains in the main part of the toolkit did you use for this report? (Technology use & development; Safety; Effectiveness; Cost-effectiveness; Organisational aspects)

No qualitative responses were sought for this question

Question 5: Can we improve on the checklists within these domains? Is the balance of questions right? (too superficial/too in-depth)

- "Section 5.1 - Technology use domain
  - 7.- Are there any differences in the use of this technology within the target setting (than the uses proposed within the HTA report for adaptation)? It is advisable to consider that answering this question could take a long time in some technologies, as could be needed to gather information.

  Section 5.2 - Safety domain
  - Many times the safety domain is evaluated "inside" the effectiveness domain in this case the questions should be formulated in a slight different way, as there is no clear safety section on methods. For example, in case we are evaluating the safety domain when we have only information about the methods on effectiveness: Was the search for studies reasonably comprehensive and is expected to include the safety aspects? In any case we suggest to include: was there a specific search to gather side effects?

  Section 5.4 - Economic evaluation domain
  - We think it is a too large section. Questions number 5 and 7 were answered on occasion of the effectiveness domain, so it could be jumped if the whole report is being adapted. Questions number 11 and 12 could be joined into a question. Questions number 23 (What equity assumptions have been made in the analysis? For example, are QALYs gained by any individual considered equal?) is difficult to understand, perhaps it could be matter for the glossary. Questions number 28: These issues are not completely related to the economic area, some of them should be under the organizational section (Technological context; Personnel characteristics; Epidemiological context (including genetic variants); Factors which influence incidence and prevalence; Demographic context; Life expectancy; Reproduction; Pre- and post-intervention care; Integration of technology in health care system; Incentives)."
- "As checklists they are all relevant and reasonably balanced. As toolkit, I don't find them especially helpful. I many respects, the toolkit is basically a checklist and a reference list. It would make more sense if the checklist and the reference/resource list were better incorporated with more direct links between them, - possibly in a software kit. If one reads and understands all the references in the resource boxes, one would not need the checklists, as the checklists represent the essence of the references."
- "We understood that this toolkit has been done for HTA on treatment. In the assessment, we had a screening HTA report where the toolkit did not suit properly. If there's just one toolkit for different kind of HTAs, the questions should be more generic. Also, differences on research questions raised problems. Our research question was somewhat different to the one in the report assessed and therefore all questions on transferability weren't relevant. The questions related to the assessment or transferability should be separated more clearly (although you have marked those a) to c). In some questions we discussed if the question are related to the assessed report or to the transferability. There were very many and quite long questions which made answering time consuming. Could a 'structured questionnaire' be more suitable?"
- "Section 5.1
  - We would answer question 2 later on in the HTA report, in the effectiveness part for example, organisational aspects… We would not present that right from the beginning but rather as the result of a full assessment.

  Section 5.3
  - In order to make it more friendly, could we gather some questions under one term or summarize them with two words so that the user knows what the questions are about. For example: Search (4,13), Selection of studies (5,6,7), Validity of studies (8,9), Findings (10, 11), Conclusion (12)

  Section 5.4
  - What about reports that only provide the findings of their literature review without presenting an original economic analysis ? Shouldn't we separate the questions that only apply to an original analysis ? If yes, maybe we should add or "adapt" some questions for those that only present the review, using some questions of the section 5.3.

  Section 5.5
  - I would include under this section : reimbursement or not. Then the question 5 of the section 5.1 would move here.

  Figure 2 : What do you mean by "utilisation"? e.g utilisation under a RCT , would that answer would be ok ? utilisation under condition coverage? Or is it about indications?"
- "The checklists are extensive/ comprehensive. Not all are needed, but they are useful when they are there."

- "Technology use & development: It is enough. OK.
  Safety (of the question box7 in page 13) :
    - We recommend to delete the 2nd aspect "the aspects that should be assessed concerning the sources of safety data are:" and to put together the questions 3, 4 and 5 within the same aspect: "the aspects that should be assessed concerning the sources of information on safety data are:"
    - question 9th Link "validity" to the Glossary and define it as internal and external validity
    - question 10th is a too in depth question. It is difficult to find a report that accomplish this issue mainly assessing the "cited studies"

  Effectiveness (of the question box9 in page 16) :
    - All boxes should be homogeneous; so, we should take into consideration box 7 and we should include the 2 aspects as in safety:
      - i)- aspects of sources of information on effectiveness data
      - ii)- quality of effectiveness assessment
    - question 1: we recommend to eliminate the 1st part " What is the research question considered?" and to modify the 2nd one as follows: "Are the research questions considered within the effectiveness and efficacy section of the HTA report relevant to your HTA question?
    - question 3: we recommend to move it after "question 4" in "section b" (reliability). In this question we are using the word "relevant" as an adjective, and also because this question is related to the sources of information.

  Cost-effectiveness (of the question box 11 in page 20):
    - question 27: It could be more accurate to write the question as follows." How generalisable and relevant are the results, the validity of the data and model, to the relevant target setting?
    - Question 24 and 25, please would you mind to rewrite them again to make them understandable?

  Organisational aspects (of the figure 2 in page 23) :
    - See the headline inside the matrix (1st column) and include (/) between Organizational aspects / dimensions.
    - What do you mean by "Inter-organizational level and Intra-organizational level"?
    - We propose to add an (*) to the headlines and explain the meaning in a footnote below the table
    - What kind of organization are you thinking in? ex. "Health insurance company, or stakeholders, or governmental or private?
- "The balance seems about right, although I thought it was a bit light on questions to assess transferability (see below). There were a few questions that I found hard to comprehend. Some of the more technical questions could do with some 'hints' or explanatory notes underneath to help people answer them. Specific questions I couldn't comprehend:
    - Technology use domains: 6. Is there any consideration of when and how technical characteristics affect outcomes?
    - Effectiveness (including efficacy) domain: 9. Was the validity of all studies referred to in the text assessed using appropriate criteria (either in selecting studies for inclusion or in analysing the studies that are cited)?
    - By 'all studies referred to in the text' do you mean all included studies?"
- "We think that both the domain checklists we went through are exahustive and that the balance of the question is right."
- "The questions on relevance, reliability and transferability in each domain go in details and helped us organize the main part of our report. Logically, there are questions that could not be applied to our research question, but this does not mean that they are unneeded in the toolkit."
- "The questions on Cost-Effectiveness compared to the other domains seem too in-depth. Since there is reference to other resources, some too special questions could be removed. Overall, at the beginning of each section a mention to the more in-epth resources could be made."
- "It was mostly OK."
- "It seems to us that this depends on who is using the checklist. That is, if the HTA team that is using the toolkit is a multidisciplinary one, where different expertises and backgrounds are involved, maybe the questions are good. On the contrary if it is, e.g. an epidemiologist who has to assess the organisational and economical aspects using the toolkit, instead of an economist or sociologist, and vice versa if it is an economist who has to assess the safety and effectiveness domain, this would imply more in-depth questions or /and further specification for each item."
- "We don't have any suggestions concerning specific questions: the balance of questions looks right."
- "The main part was too comprehensive. We mainly focus on the internal validity of the SR part of the systematic review in the adoption process. Questions on safety and effectiveness overlap – could this be simplified? The wording reliability was strange, internal validity is more familiar."

- "Safety: Q2 covered by the speedy sifting process? Q3-4: merge into one question on the comprehensiveness of the search? Q6-14 Ok though 14 relevant only for drugs. Effectiveness: Q1 covered in the speedy sifting process? Q3 covered by Q4? Q7 covered in speedy sifting?"
- "We tried to use the toolkit internally, but have decided that we need to bring in external experts to do the adaptation. Based on this our first response is that it might be too detailed, but we are currently testing it with our external experts and has asked them all to give us feedback on the toolkit."

Question 6: What additional toolkit questions and resources would help in adaptation?

- "I don't need additional toolkit questions; I rather have some sharper tools in the toolkit."
- "As mentioned above, I think there should be a bit more emphasis on transferability. Otherwise the toolkit tends to be dominated a bit by 'reliability' questions (notably the effectiveness domain), which is a bit limiting. In terms of a comment on question 14 in the effectiveness domain, 'baseline risk of patients' seems to be a bit of a narrow concept, and it is not completely clear what it means (perhaps add it to the glossary?). Perhaps the question could be broadened to include the characteristics of the patients and the setting in which they live (socio-demographic and prognostic factors)? 'Risk' also seems a bit negative, and health status might be more positive. Also, I'm not sure whether the relative risk would necessarily be the same in all settings. There may be environmental, cultural, political, societal and economic factors which may govern effects (unless I have mis-read the text). It would be useful to have further questions about the representativeness and generalisability of the trials included in the HTA report. For example, there could be further questions about the patients (e.g. whether the trials were highly selective, or whether they included a broad range of patients and sub-groups); the intervention characteristics (e.g. whether it would be feasible, and acceptable in the target population - particularly important for public health interventions); the setting and organisation (e.g. whether there is adequate capacity and structure to provide such an intervention in the target location). I came across this resource which might be useful: Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity http://www.biomedcentral.com/1471-2288/6/56/abstract Also, I noted one of the resources listed dates back to the early 1990s, and I wondered if it was still relevant? A study to assess the validity of an index of the scientific quality of research overviews, the Overview Quality Assessment Questionnaire (OQAQ). A final point is that for the most part the toolkit doesn't refer to specific kinds of interventions (e.g. drugs), but is more generic. This is a good thing as it should be ideally used for any type of intervention. However, in question box 11, question 6 it makes reference to 'evidence of the product's efficacy…'. This is probably inherited from the source material the questions are drawn from (eg Drummond; CCOHTA etc). I would recommend changing the language in this question and wherever else appropriate."
- "Would it be possible to add, in some section, a question like "What are the proposals for future research?". We think that it will be useful for the new search (when there is a need of updating the literature review)."
- "May be legal aspect can be included, with regard to how a country deal with new technologies. But it is difficult to find studies with this kind of information."
- "Additional resources could include standard result extraction sheets. This would be very helpful to develop own ones adapted to the report topic. Such extraction sheets should include the type of result to be extracted (i.e. columns on RR, NNT, ICER, etc.). This can help to point attention of extractor to the relevant parts of results reporting."

"No need to make this more comprehensive, rather aim to simplify."

Question 7: Did you use the glossary? If so, was it useful?

- "it was useful. We suggest to include economic terms, for example we would like it include a definition of 'relative cost'."
- "Consensus of the terms could help using the glossary."
- "as we are not completely familiar with the terminology it proved to be a very useful help."
- "No – too time consuming."
- "did browse it briefly and noticed a typo: 'speedy shifting' (!)"
- "Glossary was useful, but we missed some economic´s words, that we recommend to define and to include, as: "allowance for uncertainity", "stochastic sensitivity analysis". Also, we recommend to leave only one definition per term, in the Glossary."
- "The questions were clear enough for the economic evaluation domain, but probably this is because the person who was adapting this domain was always present during the construction of the glossary. We still think that it may help those who has to use the toolkit, especially in their first times. The glossary has been

used in the effectiveness domain. It appears to be clear and exhaustive, but sometimes tends to be quite lenghty; our opinion is that glossary can be improved giving at first a definition ("official definition") and an example of the term explained, then reporting comments or definitions by other single partners."

- "In some questions it helps make clear what you are looking for. The glossary is very well done. What we found useful and inspiring was that each term can be explained by different agencies, which brings a better overview of the single term. Also the direct hyperlinks in the text are very useful and time-saving."
- "Yes, specially at the beginning. It was mainly useful."
- "I used some parts of glossary. It is good to have glossary, because it helps to understand different questions the some way and properly."
- "The glossary was absolutely relevant and interesting, but was not used for the process of adoption."

Question 8: Did you consult anything other than the toolkit e.g. resources, checklists to help you adapt this report?

- "We did not consult any given resources or checklists, although those were usable and many were familiar to us. Could those be categorized better? We used WP4 EUnetHTA Core Model to see what questions ('issues') are under the topics utilization, work processes etc. Now there are some faults in the list of topics. I'll send you the last version of the organisational domain where you can see the topics and issues and correct those in your paper. Is the box of micro and meso and macro levels unnecessary? We didn't quite understand what you mean with it. Is it a checklist of the topics concluded in the report in different levels?"
- "I found the additional resources quite useful."
- "The BNF – in order to clarify the current licence indication for infliximab in the UK."
- "in our case we gave a look to "Critical appraisal worksheet for therapy" to get parameters for determining the meaning of "appropriate criteria" in the assessment of the validity of the studies considered."
- "We found very useful the list of resources in each section. It helps solve any methodological doubts and provides a wide source of extra information that can enrich the adapted report. (However, some links provide only an abstract of the papers.)"
- "I consulted with rheumatologist from Tartu University Hospital about the relevance of report in Estonian context and about treatment strategies used in Estonia. I was informed about rheumatologists interest to the content of report."
- "Yes , internal checklists for assessing systematic reviews."
- "Yes , NOKCs checklists for assessing systematic reviews."
- "We find that we need to include at least clinical experts, but perhaps also other experts depending on the content of the report and the level of ambition (inclusion of local organisational issues etc.)"

Question 9: What additional work was required to adapt this report for your target setting (your local context)?

- "There are same questions that would need much more time in case a real adaptation, like: 'Would you expect the baseline risk of patients within your own setting to be the same as the baseline risk of those patients considered within the HTA report for adaptation? (assuming that patients receive the same treatment and same comparator)'. We should have done a complete exploration about our patients`risks."
- "I had a meeting with specialists in the field to discuss the clinical relevance of the report. I quickly gathered the primary local cost data, and started at complete rebuild the economic model, as different cost and discount rates should be applies to make the model in line with Danish guidelines. I never managed to remodel the results in the HTA report; some information might be missing in the model documentation. Finally I gave up. If it had been important to complete the adapted HTA, I would probably have needed to get hold on the original model from the authors. An update of the literature review would also have been needed."
- "he report we assessed was not completely relevant for us. Only some parts of the report were usable for us.
- We would need more information about the issue from our context. For example we are interesting in the incidence of the disease in Finland."
- "We did not really adapt the report. We presented its method and findings. It was only one among several documents for our HTA report. We analysed and presented 6 guidelines and some original studies. I strongly hope that in the future we will really adapt HTA report… That leads to the next question below."
- "In the end we did not manage to adapt the report at this point of time. I have seen a problem in the DRG coding system for Slovenia as it has shown that there was only 1 laparoscopic surgery performed last year. We are now trying to analyze these data."
- "Epidemiological data: incidence and prevalence information."

- "Some brief reading of documents to assess the current need for the intervention in the UK (e.g. local epidemiology, incidence and prevalence of Crohn's disease; clinical guidelines), plus and for information on the licence of the drug (BNF)."
- "In section 5.5: data related to the organizational aspects domain. Updating the literature search related to effectiveness and safety."
- "After having answer to all the questions it is possible to have an idea of which parts can be adopted and which can be adapted. However it is still necessary a collective work, together with those who have run an analogue study in our Country (or those who have asked for the adaptation of the report) to comment on the adaptation. In an adaptation process, more than one report is appraised and we guess that some time has to be cosidered for the studies that are undergoing the speedysifting session but it is then decided not to proceed with the mail toolkit. What we want to point out is that one thing is having the idea, thanks to the toolkit, on wheather the report can be adapted to our setting, and another is really making a new report or writing the indications for polititician and decision makers. We have not finished this last process yet."
- "It was necessary to supplement the data with local context data. For the safety domain we searched official data on the legal status of the evaluated medicines in Spain. We consulted with a local neurologist the use of the four medicines in Spanish clinical practice. To evaluate the effectiveness we made a systematic revision of literature similar to the HTA report, extended to Spanish language and updated. For the cost-effectiveness analyses we looked for official data on prices and doses of the evaluated medicines."
- "We extracted results from the report and discussed it in the context of other reports on the same topic. We updated the literature searches for effectiveness and cost-effectiveness and we conducted a survey on the utilization of the technology in Germany."
- "There is a need about analyse of different treatment strategies used in Estonian context in comparison to those presented in report. We do not have any cost data about time of treatment by different specialists, about number, length and unit costs of primary care services per person using primary care by type of staff, also unit costs, number and average net ingredient costs of frequently used drugs by type of drug, about use and unit costs of other healthcare professionals and procedures per person using other health and social care, about use and unit costs of aids and adaptations. In Estonia there does not exists even several specialties (chiropodist, phlebotomist, dietician, herbalist etc.), also functions of practice nurse in primary care are probably partly different than in UK."
- "Additional work includes: economic data collection, organizational setting analysis.
- "The report need to be updated. The domain of the report regarding organisational aspects will be taken into account, verified and maybe further explored."

Question 10: Is there any other information you would like to provide us with to help improve our toolkit?

- "We don't see a clear difference between the most part of the adaptation toolkit and others checklists for reviewing literature. Sometimes we find reviews of economic studies instead of cost-analysis studies, the economic section of the toolkit is valid for cost-analysis studies and it would be needed a tool for economic reviews."
- "also I would find it helpful if the toolkit addressed the question related to economic models - how to obtain the electronic version – this might even be part of the speedy shifting section, as it often will be very difficult to adapt a HTA report without the model."
- "I have one remaining question : when there are several HTA reports on the same issue, on which criteria should I choose one among several."
- "The toolkit is extensive and a good guidance esp. for "young HTA-researchers". It is overly extensive for older researchers."
- "In some organizations, as ours, it is possible that different people assess different parts of the report, according to their background. Teresa was assessing the economic domain and Alberto the effectiveness domain. A discussion on the most relevant questions was undertaken at the end. On a cognitive point of view, the toolkit is giving many information, and going through it gives you the overall idea on adaptation. Nevertheless, it could be useful to have a space where to insert the parts that we consider as the most relevant while answering at the toolkit. As we have already mentioned, we are usually requested to adapt one report in order to answer to a particular question posed by decision makers or polititicians and time is usually an important issue. The use of the toolkit may appear too long if after having answer to, it is still necessary to write the final output. To this extent, what it could be interesting is an electronic summary of the answers we electronically give, in a form of a table so that if the toolkit is used for more than one report, it is then easy to compare results and choose which part to select from one report and which one from the other."

- "In the domain CE question 28 it is unclear what is meant by 'l. Reproduction'."
- "It will be tricky question how to adapt cost information, which is extremely interesting. How to act when price structure and consumption structure are not comparable?"
- "It could be useful: - to include among pre-criteria for selection of reports the use of a standardized index. - to suggest that each toolkit's domain would be assessed by an expert of the specific discipline."
- "With a bit of disappointment, we found that the toolkit was of very limited value. After the first reading of the report, we were already aware, at least generally, of the parts that needed adaptation, improvement and update or that could be directly adopted. The use of toolkit helped us to verify that our first opinion was based on all relevant aspects, a sort of second check. In this applicability testing round, the checklist did not modify our first judgement but highlighted parts to work on."

Additional speedy sifting question suggested by one respondent in response to Question 3

Content

1. What is the research question? Do you consider it clearly stated?
2. Are there research questions within each domain? Please list them
3. Is the health technology described? Make a brief resume
4. Are alternative technologies described and/or assessed?
5. How is the "topic" delimitated from related areas?
6. Who is the target group for the technology? – Is the target group described and delimited?
7. Who is the target group of the report? Clinicians? Clinical decision makers/administrative staff at hospitals? Politicians?
8. Has the viewpoint or perspective of the analysis been stated clearly? E.g. Is it primarily a societal perspective, third-party payer perspective, technical, economic or patient perspective?
9. Is the (regulatory) status of the technology mentioned in the report? E.g. market admission, status in other countries?
10. Is safety, harms, side effects, long term effect treated?
11. Are organisational aspects relevant? In what way?
12. What organisational level does the report refer to and what organisational aspects are treated? (see fig 2 in Adapt. toolkit*)
13. Are ethical aspects relevant (in your opinion)? Are they addressed in the report?
14. Are patient-related aspects relevant (in your opinion)? Are they addressed in the report?
15. What economic aspects are treated?
16. Is there a description of the consequences for/future impact on the health care system? Make a brief summary

Methods

17. Has the report been externally reviewed?
18. Is any conflict of interest declared? How can it affect the recommendations?
19. What methods are used to answer the research question/s in each domain? E.g. literature searches, empirical studies?
20. Are the measurements for the technological and economic domains clearly stated?
21. What are the presumptions in the analysis of ethical, patient and organisational aspects? What are the consequences for the conclusions of each domain?
22. Is the search strategy for each of the research questions described? Are comprehensive criteria for inclusion of studies stated?
23. Do the methods used for assessing the different domains answer the research question/s?
24. Is there any consideration of the quality of the literature? (Avoidance of bias) E.g. age of the data, usage of only RCTs, publication bias, peer review.
25. When was the work that underpins this report done? Does this make it out of date for your purposes?

Discussion/Conclusion

26. Are the conclusions in the report supported by the data and different analysis in the report?

Questions for you and your institution

27. Who suggested the HTA report to be transferred to your national context?
28. Who are the stakeholders inside and outside your institution regarding this HTA? – please state their possible intentions?
29. What is your overall impression of the report?
30. What additional information – if any - do you need to be able to assess the transferability/relevance of the report?
31. Considering the above – should the report be transferred? Please list the reasons

Appendix 3: Interview schedule

Region Veneto	Teresa	6th June
NOKC	Inger	13th June
HAS	Celine	HTAi Monday
OSTEBA	Rosa and Nieves	TBC
FinOHTA	Kristian	TBC

One hour interview – telephone, e-meeting or face-to-face

Interviewer's style

Detach from toolkit development – provide interviewees with this information before the interview

Balance of listening and using prompts

Interview format

To build on responses from applicability testing evaluations.

Use open questions – what did you like about it? What would make it better?

Ideas for interviews

What additional information do we want to gain from this form of evaluation? Already have evaluation sheets –

- can use it to ask respondents to expand on their suggestions and comments,
- to make suggestions and comments that are more difficult to put in writing,
- to ask for clarification,
- for the interviewer to identify issues of misunderstanding (that the respondent didn't realise they misunderstood),
- for the respondent to be more 'frank' about what they thought,
- prompts may help to identify new ideas and suggestions

Other ideas for interviews

Need for practical examples of how the toolkit can be used – take from experience of users (of toolkit when they were applicability testing). Illustrations by examples.

Ideas on making the toolkit user friendly

Ideas for transfer of HTA cost-effectiveness models

Expanding on information in the evaluation sheet - Celine

How long did it take you to use this? – 2 hours. Expand – most time spent on what? How did you use it? Looked at speedy sifting first – process of going through toolkit?

Expand on need for improvement of checklist questions – go through process of using the toolkit with Celine for these sections.

Comment – didn't really adapt the report, just used method and findings – how was toolkit helpful for this? What would you have done if the toolkit was not available. Would you want the toolkit as an aid for future reports – if so why?

Did you use the resources at all? Are these a useful addition?

Appendix 4: Responses from interviews

Why this topic?

1. Dept was thinking about rheumatoid arthritis drugs. There is a national project on these drugs in our region.
2. Proposed that they would assess 4 topics: 1 treatment, 2 services, 1 intervention. So far, have tested toolkit on one treatment and one intervention. These two topics were prioritised for our programme for 2007.
3. UK HTA programme report on PKU screening.
4. Agency interested in this technology. Started a report in this area a year ago. Hope that the final adapted report will be used by the agency.

Speedy sifting

1. Used speedy sifting in a different way – not to assess relevance across reports but relevance within the one chosen report.
2. Used speedy sifting to decide which report to take further. It was most updated, relevant and best quality. Maybe want questions in speedy sifting to ensure that the parts of the report you want to assess are within this report.
3. was very useful but self explanatory. This should have tick box answers to questions – yes, no, other etc. For this test. Two reports were identified that appeared to met requirements, decided through speedy sifting to take forward one of these.
4. Helpful, useful. Easy to follow.

Why this report?

1. Chose this report as it is a review. Very complete. Other reports were comparisons of lots of drugs. This was the only one dealing with our specific question.
2. Dental – only report in this area; Physio – 2 reports.
4. Agency interested in this technology. Started a report in this area a year ago. Hope that the final adapted report will be used by the agency.

Process of using toolkit

1. Teresa worked on the health economics part in Denmark and Alberto worked on the effectiveness part in Italy. They corresponded by phone and e-mail. Teresa explained how the toolkit worked to Alberto (not involved in the project). He is a Pharmacist. He thought the Effectiveness checklists were very good and useful. Took one week to chose the report for adaptation. Teresa and Alberto then worked on the two sections separately. Both filled out forms.
2. Had 2 people working on the reports independently. Dental – this person found the toolkit too comprehensive for her purpose. Most satisfied on speedy sifting. Physio – both speedy sifting and main part useful. Compared with own process/checklist. Safety and effectiveness sections could be merged. Time taken and process – speedy shifting 30 mins on each report. Depends on how easily information is found. 1-2 hours on main part.
3. The toolkit does not lend itself well to adapting screening reports. Either the toolkit should have i) very generic questions that apply across all HTs, or ii) generic questions plus specific questions related to health technologies (possibly in a separate section of the interactive version). At the moment, some questions are quite specific to treatments and cannot be applied to screening issues. There are too many questions within the toolkit. It took too long to complete. Organisational aspects dealt with in this UK report therefore used this section of the toolkit. This section of the toolkit needs updating – the WP4 core model has been updated. Patient aspects fall into all aspects. The aspects don't fit in all levels. Intra and Inter round wrong way. Matric very difficult to use. Not logical. Gave up! Will send grid.

There are some useful questions within the WP4 core model on organisational aspects that could be incorporated into the toolkit.

4. Firstly, we checked with other members in the agency to see if this adapted report would be useful. There was consensus to chose this one. There was only one INAHTA report in this area, so we didn't need to select which report to adapt. The language of this report was French (note – Rosa took a French report, used an English toolkit and created an adapted Spanish report!!). Both Rosa and Nieves adapted the report, French language was not a problem for Rosa but a barrier for Nieves.

Final product

1. Final adapted report would be for a Politician. Toolkit enables list of answers to questions. Language problem – translation for Politicians. Would probably require one day more work to produce a report for a Politician based on the toolkit output. Is further information required other than the product of using the toolkit? Equity aspects important here. There are criteria for determining whether a drug should be free or not.
2. short technical note/review produced. Provided info from report. Added additional info for Norweign setting. Main part of report from SR. Both reports will be used by HC managers and decision makers. These are the customers when doing the project. One for directorate of social affairs – guidelines on dental follow up. Hospital management in Norway. Strategies for older people.

If had another topic – would you use the toolkit again?

1. For Teresa, if just want to quickly make an assessment on whether to use a report it is quicker for her to just read through the report and make a judgement. But good to have something written down. Good as guidelines. Speedy sifting very useful in deciding whether to adapt. Maybe too many questions – but all useful. Difficult to chose. Experienced researchers know when to stop asking questions and make a decision on adaptation. Would use toolkit again – to remind me of everything. Not useful if just want an idea if worth adapting but important if preparing a report for a Politician.
2. Will be used. S.s. very important. Usefulness – need to translate into Norweign! Then would be adapted into organisation. Main part – possibly need more synthesised version. Then detailed for those who really need. Have been adapting HTA reports in NOKC for years therefore simplier the better
4. We would like to use the toolkit again. We like this paper version and have no need for a web based version.

Examples

1. Examples in toolkit would be useful. Happy for us to use Regione Veneto work as an example.

Explain your response to question 3

1. Is the answer to the study given? Review was assessing effectiveness and cost-effectiveness (as stated in title). But it also assessed safety. Not clear that it assessed safety if just searched using the title of the report. Need to read whole report to decide what is in it.

Glossary

1. Need for definitions for each term. Chose which definitions are most important. General very quick definition needed for each – agree between members. These need to be easily understood by people not involved in the EUnetHTA project. Alberto found the glossary confusing.
4. We found this 'messy'. Better to just have definitions in the toolkit.

Clarity on question 8 response

1. No need for inclusion of this checklist in the toolkit. Lots of repetition.

Clarity on question 9 response

1. IT tool needed. Helpful if it included reports from other users.

Clarity on question 10 response

1. Apart from glossary problem, sometimes people not involved in HTA working on adapting a report: Pharmacists, Drs, engineers, statisticians. Need definitions for words for these users rather than explanations.

Resources
3. These need to be categorised within the toolkit and we need to provide information on where these resources came from.

Recommend use to others?

1. Would recommend use by other people if electronic. Want to be able to look at reports from other users.

Cost-effectiveness

This section was not used. No cost-effectiveness analyses in this French report. But we think this section of the toolkit is very long and complex. Could you simplify some of the questions? – difficult to understand.