# Supplementary Table 5a: Extraction of methodological documents for evaluation of medical devices directly used by the institution

| **HAS (Rapid assessment method for assessing medical and surgical procedures, 2007 (18))[[1]](#endnote-1)** | **NICE (Diagnostics Assessment Programme Manual, 2011 (21))[[2]](#endnote-2)** | **NICE (Medical Technologies Evaluation Programme, 2011 (22))[[3]](#endnote-3)\*,\*\*** | **NICE (Interventional Procedures Programme Methods guide, 2007 (23))[[4]](#endnote-4)** | **CVZ (now: ZiN) (Medical tests: assessment of established medical science and medical practice, 2011, (24))[[5]](#endnote-5)\*\*\*** | **LBI (Procedural guidance for the systematic evaluation of biomarker tests, 2014 (19))[[6]](#endnote-6)** |
| --- | --- | --- | --- | --- | --- |
| **1. Assessment elements** | | | | | |
| **a) Which elements of the technology are evaluated?** | | | | | |
| Diagnostic/therapeutic benefit (safety, efficacy, contribution to treatment strategy)  Public health benefit (morbidity and mortality, quality of life, care system, public health policies and programmes) | Test accuracy, clinical effectiveness, cost-effectiveness | Efficacy, effectiveness, usability, safety outcomes (including intermediate clinical outcomes) and available clinical and health economics studies of any type | Efficacy, safety | Analytical accuracy, diagnostic accuracy, clinical utility (test-plus-treatment-strategy), cost-effectiveness | Clinical utility, clinical management effects due to testing, direct health effects of testing, emotional, social, cognitive, behavioural responses to testing, legal and ethical effects of testing |
| **b) Are there methodological approaches specified for each aspect/element?** | | | | | |
| No | Yes, for effectiveness and cost-effectiveness | No | Yes, for efficacy and safety | Yes, for clinical utility | Yes, see above |
| **2. Evidence procurement and selection** | | | | | |
| **a) Is the assessment based on submissions or on internally conducted research?** | | | | | |
| Analysis of identified published data, issuing of an opinion by a working group of professionals. | Diagnostic Assessment Report produced by an External Assessment Group following the DAP Manual methods: contains a systematic review of the clinical and cost effectiveness evidence and cost effectiveness analysis.  No formal manufacturer submission. | Submission from the sponsor (clinical and economic evidence).  Assessment report presented by the External Assessment Centre (independent of NICE; analysis and appraisal of submission; e.g. reproduces the sponsor’s search to validate that all relevant evidence has been identified).  Further evidence (e.g. from the Programme team) and contributions (e.g. from expert adviser). | Primarily on systematic literature search (done by the Institute’s Information Services team and Programme’s technical team).  Programme team prepares evidence for each procedure, which forms the basis of the Committee’s decision making.  Commentary from Specialist Advisers or consultees. Evidence from manufacturers not routinely used. | Applicant expected to submit data on which basis ZiN can form an opinion.  A systematic literature search is also carried out.  Consults always relevant experts. | Systematic literature search (for more details, see internal manual of LBI) |
| **b) What types of evidence are admissible for the assessment?** | | | | | |
| Based on following level of scientific evidence (HAS grading scheme):  I- High-powered RCT', meta-analyses, decision analyses. II- Low powered RCTs, non-randomised trials, cohort studies. III- Case-control studies. IV- Retrospective studies, case series, descriptive epidemiological studies, controlled trials with bias. | Test accuracy: prospective cohort, cross-sectional, retrospective case-control studies.  Test side effects: high-quality systematic reviews, RCTs and other comparative designs, cross-sectional, case studies, patient registries Existing models: may be used if appropriate and of high-quality for management and treatment following a diagnosis Treatment effectiveness: high-quality systematic reviews, individual RCT or a meta- analysis, other comparative designs  Care management: Clinical guidelines from NICE and other organisations, diagnostic before- and-after studies, expert clinical input  Manual acknowledges that end to end studies are rarely available for diagnostic technologies and encourages inclusion of all evidence types and the application of a linked evidence approach where appropriate. | Primary clinical research or secondary research (such as evidence synthesis, network meta-analysis, modelling studies). | Study design: meta-analysis, RCTs, non-randomised controlled trials and case series studies the main sources of data if RCTs not available. Study size: priority for highest number of participants Follow-Up: priority for longer and more complete follow up | Direct Evidence: RCTs of high-quality, sufficiently long duration. Non-randomised comparative study, cohort study or similar may be accepted in the absence of RCTs.  Indirect Evidence: Application of a comparative analysis framework. | The most appropriate study design to evaluate the impact of a bio-marker test on clinical management effects and health outcomes is a RCT.  The assessment of biomarkers does not necessarily include a classical intervention research question, but instead may include questions on prognosis or diagnostic accuracy, depending on the context of use. For some of these questions the only evidence feasible and/or ethical will be from observational studies and different evidence hierarchies may apply: e.g. from NHMRC and the Oxford Centre for EBM, which should be considered when prioritising available evidence.  Diagnostic accuracy studies are cross-sectional by nature. Study designs are differentiated in ‘single-gate’ (diagnostic cohort study) and ‘two-gates’ (diagnostic case-control studies) studies (Table 5.2-1, p.24 for the Evidence hierarchies by research questions). |
| **c) What sources are used to identify the evidence (databases, registries, specific websites, grey literature)?** | | | | | |
| Medical and scientific literature databases published national and international clinical practice guidelines, consensus conferences, systematic reviews, meta-analyses and other assessments; relevant websites (government agencies, learned societies etc.); grey literature and relevant legislative and regulatory texts | Literature databases and specialist databases (e.g. MEDION, ARIF, DARE), unpublished data (from manufacturers and experts), bibliographies of identified studies, conference abstracts | Published evidence: A range of medical literature databases, including primary research databases; registers or databases of systematic reviews; meta-analyses and technology assessment evaluations; registers or databases of ongoing clinical trials; and conference proceedings.  Unpublished evidence sources: as part of submission; identified by External Assessment Centre (e.g. professional or manufacturer-sponsored registers)  Contributions from expert adviser, patient groups. | Published and unpublished evidence  Sources for evidence from systematic reviews and health technology assessments (e.g. DARE), primary research evidence (e.g. CENTRAL), ongoing research databases (e.g. National Research Register), conference proceedings | No information | Challenges associated with literature search for medical tests:  1. Due to still underdeveloped indexing and reporting of studies of diagnostic tests, literature search should not rely (exclusively) on diagnostic search filters, in particular these filters are inappropriate for systematic review of clinical effectiveness.  2. If the name(s) of the diagnostic test(s) relevant for the research question is not known, search strategies should capture the ‘concept of diagnostic tests’  3. To identify all studies for a systematic review, searches should include text words (not subject headings alone) and be combined with hand search including additional sources of information: specialised databases, citation tracking and regulatory documents). |
| **d) Are there specific requirements as to the endpoints included for assessment?** | | | | | |
| No information | All health outcomes resulting directly or indirectly from the use of the test, informational outcomes of value to the patient. Longer-term outcomes in most cases. All costs stemming from the use of the test should be included. | Analysis of indirect and intermediate clinical and system outcomes.  If not available in full, appropriate modelling of outcomes may be submitted and should be reflected in the cost analysis. | Patient-focused outcomes (survival, morbidity, quality of life) prioritized over surrogates. Safety outcomes (SAE, etc.) to be included and searched for. | Intrinsic value of medical tests and their consequences for the health of patients.  Test-plus-treatment-strategy expressed in health outcomes and data on the analytical and diagnostic accuracy of the test. | Accurate diagnosis is a prerequisite for a successful therapy, but it should not be seen in isolation.  Benefit to patients resulting from diagnosis should be measured in patient-relevant outcomes, such as survival (mortality), clinical events, adverse events, patient-reported outcomes (health related quality of life), activity and function.  Based on the analytical framework developed, the reviewer should first explore all outcomes resulting from embedment of the test in the testing and treatment strategy in comparison to clinical practice without the test.  Reviewers should then make a careful selection of the relevant outcomes both to the process of testing and to the results of the test by mapping them according to the following categories: Clinical management effects due to testing; Direct health effects of testing; Emotional, social, cognitive, behavioral responses to testing; Legal and ethical effects of testing.  A decision which outcomes are relevant for a review depends on the type of test under review and on the needs of the stakeholders of the study. As a consequence the prioritization of the relevant outcomes should involve the commissioner of the study.  The outcomes are decisive only if they differ between current and new testing and treatment strategy.  The outcomes should explicitly be rated by importance a priori. |
| **e) Are there specifications as to the comparators to be included?** | | | | | |
| No information | The comparator or comparators are the technologies or tests that are most commonly used or are recommended in current NICE guidance for the functions in the evaluation. There may be multiple tests or variants or test sequences in common use and all are included as comparators. | The standard intervention against which the technology under evaluation is compared. The comparator is usually a similar or equivalent technology used as part of current management. The comparator can be no intervention. | The preferred comparator is the best standard treatment (rarely available for emerging interventional procedures). Depending on the circumstances, either active treatment or sham (placebo) is the preferred standard in assessing the efficacy and safety of a procedure.  Sometimes all of the evidence for a given procedure relates to non-comparative studies. In these circumstances, selected evidence of likely comparator procedures may be presented. No available comparators: key aspects of natural history (such as survival rates at different time intervals if untreated) may be presented. | Comparative test-plus-treatment-strategy (the current best/usual strategy).  All advantageous and disadvantageous differences between both strategies determine the efficacy of the new test. | Comparator in reviews of medical tests is the current test and treatment strategy, i.e. the sequence of tests, patient management decisions and treatments.  In diagnostic accuracy studies, the comparison of a new and an existing test may be direct: in fully paired direct comparisons, where all participants receive the index test and one or more comparator tests. As an alternative participants may be randomly allocated to receive the index or the comparator test.  The comparison is indirect, if the estimates of diagnostic accuracy from different study populations are compared. |
| **3. Evidence Appraisal** | | | | | |
| **Is the evidence critically appraised before conclusions are drawn? If so, are there specific tools used?** | | | | | |
| Yes, each article selected is reviewed according to the principles of critical appraisal of the literature using checklists. A level of scientific evidence (HAS grading scheme) is allocated to each study. | Yes, the critical appraisal is done by the External Assessment Group (EAG) when preparing the Diagnostics Assessment report.  DAP Programme Manual recommends the use of QUADAS (or a modified version). | Yes; a detailed analysis and critical appraisal of the submission in the form of an assessment report is done by External Assessment Center. Guidance document includes Committee’s recommendations and considerations, including key evidence taken into account by the Committee, its view of this evidence, and the areas of contention and uncertainty that arose during the discussions as well as contributions from expert advisers and patient and carer organisations. | Yes; a critical appraisal highlights specific methodological issues or concerns about the evidence (e.g. study type, quality of study; effect size), as perceived by the Programme’s technical team.  No checklist used (relative importance of the dimensions will vary according to the intervention).  A list of issues and evaluated aspects is specified in the document. | QUADAS (or modified version)  Gradually starting to use the GRADE method in assessing tests | Standard internal validity critical appraisal checklists for RCTs and cohort studies for interventions may be used (see internal manual of LBI).  Diagnostic accuracy studies: QUADAS-2 checklist  Prognosis studies: QUIPS  GRADE incl. adaptions of GRADE framework to diagnostic and prognosis studies  Framework for evaluating evidence on the clinical benefit of co-dependent technologies for reimbursement decisions |
| **4. Review Process & Transparency** | | | | | |
| **a) Are stakeholders directly involved in the production process?** | | | | | |
| Yes: Working group of health professionals from a number of disciplines, working in different types of public or private practice from all over the country and from all schools of thought meets once and systematically discusses whether the literature review and their own experience meet the criteria used to measure the procedure's clinical benefit.  If possible, consensus is reached and a report produced. | Yes: Stakeholders are invited to a scoping workshop, to comment on the Diagnostics Assessment Report produced by the External Assessment Group and to take part in the public consultation on the draft guidance recommendations.  Specialist Committee Members are recruited to the Diagnostics Advisory Committee for each topic; these are people who are clinical specialists in the area or patient experts. | Yes: Expert advisers contribute to the evaluation of technologies by providing additional knowledge, opinion and experience to the Committee.  Expert advice can also be used as part of evidence synthesis or modelling studies.  The Patient and Public Involvement Programme (PPIP) always approach patient and carer organisations to obtain their views on the technology. | Yes: Using the expertise and knowledge of Specialist Advisers. Involving the notifier of the procedure, manufacturers.  Commentary on patients’ experiences of the procedure is considered by the Committee in formulating its final recommendations. | Yes: Relevant experts are consulted, depending on the subject.  Where necessary, methodological experts will be included, in addition to the experts on the subject matter. | No information, but the document is meant to complement the internal manual of LBI. Thus information might be given there. |
| **b) Are (other) specific measures taken to ensure transparency of results?** | | | | | |
| Summary and the complete report available online along with English translation of the summary. | Draft guidance is produced after committee decision and is made available for public consultation for 4 weeks.  Specialist Committee members need to declare conflicts of interests.  Guidances are published on the website of NICE. | Draft guidance is produced after committee decision and is made available for public consultation for 4 weeks.  Specialist Advisers are required to declare any conflicts of interest.  Guidances are published on the website of NICE. | Draft guidance is produced after committee decision and is made available for public consultation for 4 weeks.  Specialist Advisers are required to declare any conflicts of interest.  Guidances are published on the website of NICE. | Only found: In order to promote the quality of assessments, it is particularly important to ensure that assessments take place transparently and, where necessary, to seek the allegiance of external experts on the subject matter and methodology. | No information, but the document is meant to complement the internal manual of LBI. Thus information might be given there. |
| **5. Reassessment** | | | | | |
| **Are there specific intervals/ conditions set for reassessment (fixed time/ new indication)?** | | | | | |
| Only found: 'All searches are updated until the end of the project.’ | The Diagnostics Assessment Report will be updated if it were to be decided that an update to the guidance is required. | Programme team updates the literature search every 3 years to ensure that relevant new evidence is identified.  NICE may review the guidance before the expected review date when it becomes aware that there is significant new information that it considers is likely to have a material effect on the recommendations in the existing guidance.  MTEP team may recommend to Guidance Executive that guidance be withdrawn. | During the consultation period, an updated literature search is conducted to find any further evidence that may have been published since the first Committee meeting.  For some procedures, NICE will become aware that potentially important research is due to be published, maintains active surveillance of these procedures and brings a revised overview to the Advisory Committee when appropriate. | No information | No information, but the document is meant to complement the internal manual of LBI. Thus information might be given there. |
| **6. Knowledge Exchange and Transferability** | | | | | |
| **a) Does the agency use data and/or reports from other HTA institutions?** | | | | | |
| No information | No information | Databases of systematic reviews; meta-analyses and technology assessment evaluations were searched. | Evidence from systematic reviews and health technology assessments (e.g. DARE) | No information | No information, but the document is meant to complement the internal manual of LBI. Thus information might be given there. |
| **b) Are there specific processes in place to ensure transferability of results?** | | | | | |
| No information | Differences between study and application context (e.g. patient population, intervention setting, locality) should be documented, with reference to the decisions taken about which studies to include in the assessment. | No information | No information | Only found: Following the steps of EbM | Use of the PICO scheme |
| **7. Cost and Economic Evaluation** | | | | | |
| **a) If cost or cost-effectiveness information is considered, what methods of evaluation are allowed?** | | | | | |
| No information | For the reference case, cost-effectiveness (specifically cost–utility) analysis is the preferred form of economic evaluation. This seeks to establish whether differences in costs between options can be justified in terms of health effects related to quality of life. | The approach expected to be appropriate for most technologies is cost-consequence analysis.  Such analysis may not be needed if relevant high-quality economic evaluations are already available. | Cost effectiveness not assessed | No information | Methods of cost-effectiveness excluded in this report. |
| **b) If cost or cost-effectiveness information is considered, what types of analyses are allowed/recommended?** | | | | | |
| No information | The DAP considers cost effectiveness analysis with health outcomes expressed as QALYs. | Cost-consequence analysis considers the costs and resource consequences resulting from, the use of the technology under evaluation and comparator technologies, as well as considering relevant clinical benefits alongside the cost analysis. | Cost effectiveness not assessed | No information | Methods of cost-effectiveness excluded in this report. |
| **c) If cost-effectiveness information is considered, what kind of perspective must be used for the cost-effectiveness assessment?** | | | | | |
| No information | Reference case: financial costs relevant to the NHS/PSS. As far as possible, estimates of unit costs and prices for particular resources should be used consistently across evaluations. Diagnostic tests should generally be priced at average cost. | Models should capture and quantify the impact of introducing a new technology into current healthcare pathways and routine NHS use. | Cost effectiveness not assessed | No information | Methods of cost-effectiveness excluded in this report. |
| **d) If cost-effectiveness information is considered, is estimation of uncertainty mandatory (e.g. confidence intervals)?** | | | | | |
| No information | Sensitivity analysis should be used to explore uncertainty around the key structural assumptions used in the analysis (scenarios and parameters). | Uncertainty analysis techniques (relating to chance, evidential and model uncertainty) should be undertaken (e.g. scenario-based deterministic sensitivity analysis, threshold analysis or probabilistic sensitivity analysis) | Cost effectiveness not assessed | No information | Methods of cost-effectiveness excluded in this report. |
| **e) If cost-effectiveness information is considered, are threshold used to determine whether technology is cost-effective?** | | | | | |
| No information | Expected net monetary or health benefits may be presented (in addition to ICERs), using values of £20,000 and £30,000 for a QALY gained. | Discounting principles consistent with other NICE guidance programmes, discount rate of 3.5%. | Cost effectiveness not assessed | No information | Methods of cost-effectiveness excluded in this report. |
| **8. Other aspects** | | | | | |
| - | Characteristics of diagnostic technologies pointed out:  Evidence about patient outcomes for diagnostic technologies is lower in quantity and quality than evidence for pharmaceutical products.  Diagnostic technologies, particularly those based on electronics, often change rapidly as new methods, upgrades and capabilities are added. | Characteristics of medical technologies pointed out in method guide: Medical technologies are different from other medical interventions because technologies may be modified over time in ways that change their effectiveness. The clinical outcomes resulting from the use of technologies often depend on the training, competence and experience of the user (sometimes referred to as the ‘learning curve’). (*continues in next page*) (…)The Committee may make recommendations for use of the technology in specific circumstances only (e.g. by staff with certain training) and may also recommend development of further evidence or the use in a research context. | The Committee makes recommendations about the procedure on the basis of the available evidence relating to its efficacy and safety. These maybe affected by certain variables about which published evidence provides little or no helpful information. The individual operator and the different devices used to perform procedures are often important in this context. Different recommendations possible. | Give details of constructing a comparative analysis framework. | - |

1. \*Medtech Innovation Briefings (MIBs): integrated process statement on how NICE select topics and prepare the documents (available [here](https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-advice/Medtech-innovation-briefings/MIB-interim-process-methods-statement.pdf)); belongs to the MTEP programme but was not considered.

   \*\*Interim addendum to replace existing section 8, Guidance reviews, in MTEP process guide was considered in the analysis; Available [here](https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-medical-technologies/MTEP-Interim-addendum-guidance-review.pdf)

   \*\*\*Update of ‘Assessment of established medical science and medical practice, 2007’ (2015, English summary available [here](https://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/publications-in-english/2015/1501-assessment-of-%E2%80%98established-medical-science-and-medical-practice%E2%80%99-update-2015/1501-assessment-of-%E2%80%98established-medical-science-and-medical-practice%E2%80%99-update-2015/Assessment+of+%E2%80%98established+medical+science+and+medical+practice%E2%80%99+%28update+2015%29.pdf)) also touched the assessment of medical tests, thus we considered the updated information given in the summary in our extraction and analysis.

   Focuses on benefit assessments carried out by HAS, specifically regarding the diagnostic or therapeutic benefit as well as the public health benefit. [↑](#endnote-ref-1)
2. Evaluates diagnostic technologies that have the potential to improve health outcomes but whose introduction is likely to be associated with an overall increase in cost to the NHS. [↑](#endnote-ref-2)
3. Selects, routes and evaluates new or innovative medical technologies (including devices and diagnostics); evaluates cost saving diagnostics. [↑](#endnote-ref-3)
4. Assessment of the efficacy and safety of interventional procedures carried out by the Institute’s Interventional Procedures (IP) Programme; mostly investigates new procedures. [↑](#endnote-ref-4)
5. The report specifically contains an elaboration of the method used by ZiN for assessing medical tests on the basis of principles of EbM. Additionally it aims to inform other parties about the principles and methods used by ZiN. [↑](#endnote-ref-5)
6. The document is meant to complement the LBI internal manual for qualitative evidence synthesis on effectiveness and safety of biomarker tests. Only steps in the assessment which deviate or need special attention are covered, and methods of quantitative evidence synthesis and cost-effectiveness are excluded. [↑](#endnote-ref-6)