NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Briefing paper for methods review working party on companion diagnostics

The briefing paper is intended to provide a brief summary of the issues that are proposed for discussion by the Methods Review Working Party to inform an update to the Institute's Guide to Methods of Technology Appraisal. It is not intended to reflect a comprehensive or systematic review of the literature. The views presented in this paper are those of the authors and do not reflect the views of the Institute.

1 Review of the 'Guide to Methods of Technology Appraisal'

The Institute is reviewing the 'Guide to the methods of technology appraisal', which underpins the technology appraisal programme.

The original Methods Guide was published in February 2001, and a revised version was published in 2007. The Methods Guide provides an overview of the principles and methods used by the Institute in assessing health technologies. It is a guide for all organisations considering submitting evidence to the technology appraisal programme and describes appraisal methodology.

The revised draft of the Methods Guide will be available for a 3-month public consultation, expected to begin in May 2011. We encourage all interested parties to take part in this consultation.

2 Background

2.1 What are companion diagnostics?

Companion diagnostics are tests that are typically developed to select patients who will benefit from specific treatments, usually pharmaceuticals, by improving the responder rates or decreasing side effects. The US FDA definition requires that the companion diagnostic provide "information that is essential for the safe and effective use of a corresponding therapeutic product". Most companion diagnostics use genetic or protein markers to identify patients who will benefit from targeted treatments. These markers to be measured by companion diagnostic tests are usually referred to in the marketing authorisation for the treatment. Examples of treatments based on specific markers appraised to date are shown in the table below.

Appraisal Title	**Marker**
Trastuzumab for the adjuvant treatment of early-stage HER2-positive breast cancer (TA107)	HER-2 (protein marker)
Bevacizumab and cetuximab for metastatic colorectal cancer (TA118)	EGFR (protein marker)
Cetuximab for the first-line treatment of metastatic colorectal cancer (TA176)	KRAS (genetic marker)
Gefitinib for the first-line treatment of locally advanced or metastatic non-small-cell lung cancer (TA192)	EGFR TK mutations (genetic marker)
Trastuzumab for the treatment of HER2-positive metastatic gastric cancer (TA208)	HER-2 (protein marker)

2.2 Regulatory requirements

The US FDA draft guidance on companion diagnostics generally requires that the companion diagnostic and the treatment be evaluated contemporaneously, although there are a number of exceptions. To date, the EMA does not explicitly deal with the evaluation of companion diagnostics. Diagnostics are regulated in accordance with the European In-Vitro Diagnostics Directive. Marketing authorisations granted for a pharmaceutical by the EMA may specify a patient sub-population requiring the testing for a

genetic or protein marker but the specific companion diagnostic to be used is not stated. In some cases the pharmaceutical SPC may indicate that only validated tests should be used.

2.3 Relevance of the topic to NICE technology appraisals

Increasingly, the marketing authorisations for new pharmaceuticals require the use of companion diagnostics. It is therefore important that within the appraisal of pharmaceuticals, adequate consideration is given to companion diagnostics. This should be balanced against the need to develop appraisals of pharmaceuticals with companion diagnostics within the normal resources and timeframes of the technology appraisals programme. NICE methods for the evaluation of companion diagnostics will develop over time and are likely to involve the technology appraisals and diagnostics assessment programmes. This review of the 2008 Technology Appraisals Methods Guide is an important opportunity to ensure adequate provision for the evaluation of pharmaceuticals requiring the use of companion diagnostic products.

The establishment of the new diagnostics assessment programme (DAP) has raised the profile of NICE with the diagnostics community and there is an expectation that NICE will evaluate companion diagnostics in conjunction with assessments of pharmaceuticals. The programme used to evaluate the diagnostic technology could be either TA or DAP depending on the question being considered. This briefing paper highlights key issues related to companion diagnostics that need consideration in the Methods Guide review.

2.4 Companion diagnostics in DAP or TA

When the marketing authorisation of a newly licensed drug includes the use of a diagnostic test to identify the eligible population the Appraisal Committee is likely to need to take the companion diagnostic into consideration when developing the guidance for the new drug. It would generally be inefficient to split the NICE processes between TA and DAP, and also this would not lead to timely guidance for the new drug. Taking account of the specific companion diagnostic used in clinical trials is also relatively straightforward as the patient outcomes observed in the trials are those from the treatment informed by that

specific companion diagnostic. Assessment of the pharmaceutical and companion diagnostic "package" can be undertaken in much the same way as for pharmaceuticals without companion diagnostics. However, in circumstances where alternative tests are available (e.g. proprietary test kits or "in-house tests" for the same marker that would fulfil the requirements of the pharmaceutical marketing authorisation), the amount of extra effort to fully evaluate these alternative options is likely to exceed the available resources and timeframe in technology appraisals.

When, after a drug is in established use, a diagnostic technology is introduced as a companion diagnostic to improve the responder rates, or decrease side effects, the diagnostic technology would typically by evaluated by either the Medical Technologies Evaluation programme (MTEP) or DAP rather than Technology Appraisals.

Companion diagnostics in Technology Appraisals

The 2008 Technology Appraisals Methods Guide refers to companion diagnostics in section 5.7.5 which reads:

"If the use of the technology is conditional on the outcome of a diagnostic test, the accuracy of the test and associated costs should be incorporated into the assessments of clinical and cost effectiveness."

In the 134 Technology Appraisals published since 2006 a specific diagnostic tool was described as part of the marketing authorisation and in the actual NICE recommendations of 47 Appraisals. Of these, the majority related to tools to assess disease severity, many included imaging, histology or other tests, and only the 5 listed in Table 1 could be referred to as true companion diagnostics.

The issues in previous appraisals around companion diagnostics were as follows:

1. Target population is a *post hoc* subgroup; The 2008 Technology Appraisals Methods Guide states: "The characteristics of patients in the

subgroup should be clearly defined and should preferably be identified on the basis of an *a priori* expectation of differential clinical or cost effectiveness due to known, biologically plausible mechanisms, social characteristics or other clearly justified factors." Often the information related to subgroups with a specific biomarker was not prospectively included in the trial.

2. Comparator data for a different population: If the data on the comparator technology are not from the clinical trial of the new pharmaceutical, then the comparator data will not usually be available for the specific target population.

3. Uncertainty over the use of the test in practice: Committee decisions were informed by clinical specialists' opinion, rather than firm evidence as to how the testing will be handled in clinical practice.

4. Test accuracy: The biggest issue relates to tests other than the specific one used in the clinical trial which may still fulfil the requirements of the marketing authorisation (e.g. alternative proprietary tests or "in-house tests" for the same marker). Often there is no evidence of the accuracy of the alternative test or its impact on the efficacy of the treatment. Tests may have serious false positive or negative rates impacting the value of testing/treatment. A second issue relates to changes over time of the knowledge base of what mutations are affected by the treatment. As more relevant mutations are discovered, the utility of any diagnostic test may change.

5. Testing increases costs for the NHS: The costs for testing all potentially eligible patients are included, but only those patients who get treated will benefit. Often the prevalence of the biomarker[1] is not known. A low prevalence of the biomarker means that more people are tested per patient identified to benefit from the new treatment which increases the cost per patient found and impacts cost effectiveness.

[1] In this paper the term "biomarker" is used in its general sense to include any biological marker that may affect the treatment. These can include nearly any lab result and is not restricted to protein markers.

The 2008 Technology Appraisals Methods Guide also includes some general coverage of diagnostics (see Appendix). Following the establishment of the DAP, standalone diagnostic technologies will not be assessed in TA and the relevance of these sections should be reviewed.

Companion diagnostics in the DAP

The DAP methods, designed for the assessment of diagnostics generally, are suitable for the assessment of multiple companion diagnostic options. They are also suitable for assessing diagnostics technologies with the potential to be used to improve the targeting or use of pharmaceuticals already used in clinical practice.

3 Proposed issues for discussion

It is expected that only the single companion diagnostic test option used in clinical trials would be fully considered in technology appraisals of pharmaceuticals since evaluating multiple diagnostic options would dramatically increase the time and resources required for the pharmaceutical evaluation. It is important, however, to acknowledge that other tests could potentially be used in clinical practice and that in using alterative tests, there is a risk that the alternative tests do not select exactly the same population as the test originally used in the clinical trials. Correspondingly different outcomes from treatment could also result. Management of this key issue within the technology appraisal of pharmaceuticals with companion diagnostics is important in ensuring optimal and cost effective use of the pharmaceuticals. This issue is avoided when the test used in clinical trials and considered within the technology appraisal is also adopted in clinical practice. In some cases, it may be possible to report the diagnostic accuracy of the test used in the clinical trials. Any alternative tests should then be validated and compared to the companion used in the trials prior to adoption. In many cases, however, diagnostic accuracy data (in this case, accuracy may mean the test's ability to predict treatment efficacy) may not be available – the only data available may be the trial outcomes resulting from treatment informed by

the test used in clinical trials. The specific test used in the clinical trial then becomes the “reference standard” with which alternatives should be compared. That is, the accuracy of alternatives tests is based on their agreement with the reference standard.

A challenge in the technology appraisal of pharmaceuticals with companion diagnostics is providing appropriate guidance and warnings on the potential use of alternative tests without detailed evaluation of the various test options. This could be as simple as a discussion within the committee considerations section or where appropriate, guidance on the diagnostics accuracy that would need to be demonstrated prior to the adoption of an alternative test. In particularly complex cases it may be appropriate to undertake a DAP assessment of the alternative companion diagnostic options following the initial technology appraisal.

A further key issue for the assessment of pharmaceuticals with companion diagnostics is how to handle the costs associated with the companion diagnostic testing. Even for pharmaceuticals that do not have companion diagnostics, the identification of patient populations for treatment often still requires significant diagnosis – and such costs are not normally included within the assessment.

Issue 1 – Should the costs of the companion diagnostic be included as part of the total costs in a technology appraisal of the treatment, and, if so, how does this impact the assessment?

Most technology appraisals start with an identified population that has been diagnosed. In this setting, the costs of the diagnostic process are not included and the diagnostic processes are generally assumed to be cost effective. The costs assessed usually begin with the treatment and include the costs of the treatment plus any further health costs influenced by the treatment or the disease in question. These can include costs of the disease and its further treatment as well as costs of dealing with the side effects stemming from the treatment or downstream treatments. The Diagnostics Assessment Programme, when assessing diagnostic tests, includes all costs stemming

from the point of the diagnostic test. The assumption is that the treatment and comparator are all cost effective.

It has been argued that one can differentiate between diagnostic processes that are carried out to diagnose a condition in general (and then choose from a number of established treatment options) and a diagnostic test that is carried out to make a decision for treatment with a specific drug. On that basis, it has been suggested that, when evaluating a treatment that has a companion diagnostic, the costs of testing should be included in the assessment. This is because the treatment cannot be initiated without the companion diagnostic and hence the cost of testing is part of the cost of treatment. However, as mentioned above, all treatments require some type of diagnosis before use, but the diagnostic costs are not generally included in appraisals of treatments.

For discussion:

1. Is it reasonable to include the diagnostic costs when looking at treatments with companion diagnostics, but not when treatments use diagnostic tests that are already is commonly in use?

2. If diagnostic costs are included in the appraisal, should it be required that separate ICERs be provided for the therapeutic with those diagnostic costs included and excluded?

3. If it is decided that costs and any direct outcomes for a companion diagnostic need to be included, what should be done when a further drug requiring the same particular companion test is subsequently appraised?

4. How should the situation be handled where the companion test is initially (but perhaps only initially) made "free" by the manufacturer?

Issue 2 – If a treatment is appraised that has been trialled with a particular companion diagnostic, what should the guidance say about the characteristics of the diagnostic test?

In most cases, information on the companion diagnostic that was used to select patients for the clinical trial(s) of the related treatment will be available. The diagnostic test will assess some marker (genetic, protein, or other) presumed to be relevant to the treatment efficacy. In some cases there will not be any other “gold standard” reference test available. However, it may be the case that the diagnostic test does not assess the marker perfectly and this may not be known. It also may not be known whether the treatment would be more effective if the test were perfect (i.e. 100% sensitive and 100% specific for the marker).

When alternative tests are available or likely to be available and used, then the question of relative accuracy becomes an issue. If there are trials of the treatment using the alternative test, then again those data would provide end outcomes directly and test accuracy, *per se*, is not an issue. If an alternative test is only compared to the test used in the trials and does not perfectly agree with that test in all cases, then there can be uncertainty about which test is more effective in maximising the benefits from the treatment.

For discussion:

1. Are there circumstances where it would be appropriate to recommend only the specific test used in the clinical trials even if this is not specified in the marketing authorisation?

2. If a true gold standard exists for the marker that has never been trialled with the treatment, under what circumstances can it be assumed that it is the appropriate marker for maximising treatment benefits? Where such a gold standard exists, should test accuracy standards (sensitivity and specificity) for alternative companion diagnostics be provided in the guidance?

3. If no such gold standard exists, should test accuracy standards relative to the companion diagnostic used in the clinical trial be provided in the guidance?

4. Alternatively, should general warnings be given on the potential consequences of using alternative companion diagnostics in the recommendations and/or committee considerations?

5. What information on the companion diagnostic used in the clinical trial(s) and the potential alternative tests should be requested as part of the manufacturer submission?

Issue 3 – Should the current sections on methods for assessing diagnostics continue to be included in the Technology Appraisals Methods Guide?

As TA will no longer appraise standalone diagnostics since those would be evaluated by MTEP or DAP, it may be appropriate to delete the current wording about diagnostics (see Appendix). A new section on companion diagnostics will probably be needed following consideration of the issues raised in this paper.

For discussion:

1. Should the current sections on diagnostics be deleted or replaced with a reference to the DAP programme manual?

4 References

FDA, Draft Guidance for Industry and Food and Drug Administration Staff - In Vitro Companion Diagnostic Devices,http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm262292.htm

NICE (December 2011) Diagnostics Assessment Programme Manual.

5 Author/s

Prepared by Hanan Bell, Nick Crabb, Elisabeth George

6 Appendix

General coverage of diagnostics in 2008 Technology Appraisals Methods Guide states

5.17 Diagnostic technologies can be used in different ways (for example, for disease identification, monitoring of disease progression and treatment, assessment of disease prognosis, or initial screening) and this should be reflected in the evidence submitted to the Institute.

5.18 Evidence for the appraisal of diagnostic technologies should normally incorporate evidence on the accuracy of the diagnostic technology. It is also important to incorporate the predicted changes in health outcomes and costs as a result of treatment decisions based on the test result.

5.1.9 The general principles guiding the assessment of the clinical and cost effectiveness of diagnostic technologies should be the same as for other technologies. However, particular consideration of the methods of analysis may be required, especially in relation to evidence synthesis. Evidence for the effectiveness of diagnostic technologies should include the costs and outcomes for people whose test results lead to an incorrect diagnosis as well as those who are correctly diagnosed.

5.1.10 As for other technologies, RCTs have the potential to capture the pathway of care involving diagnostic technologies, but their feasibility and availability may be limited. Other study designs should be assessed on the basis of their fitness for purpose, taking into consideration the aim of the study (for example, to evaluate outcomes, or to evaluate sensitivity and specificity) and the purpose of the diagnostic technology.

5.3.3 Assessments of diagnostic technologies should follow the general principles of systematic reviews as recommended here for other healthcare technologies. However, it is recognised that the specifics of, for example, the meta-analysis of studies of the sensitivity and specificity of diagnostic tests are different from reviews of the effects of therapeutic interventions. This is an area of active methodological research.