CONTEXT AND POLICY ISSUES

Oncotype DX (ODX) is a gene expression profiling test designed to measure the 10-year risk of tumor recurrence in early breast cancer following initial diagnosis.\textsuperscript{1,2} The risk of tumor recurrence is reported as a 21-gene signature or recurrence score (RS) on a scale of 0-100;\textsuperscript{2} this RS is then translated into one of three categories of risk: low (RS<18), intermediate (RS 18-30), or high (RS>30).\textsuperscript{1} In the adjuvant setting, results from gene expression profiling tests aim to complement traditional prognostic information obtained from clinicopathologic findings by adding better risk stratification, which may then assist in the identification of patients most likely to benefit from adjuvant chemotherapy.

Initially developed in women with estrogen receptor-positive (ER+) and lymph node-negative (LN-) early invasive breast cancer, the use of ODX has since expanded into other settings, including the lymph node-positive (LN+) population.\textsuperscript{3} Because invasive breast cancer is LN- at diagnosis in 65% of women\textsuperscript{2} and is where the bulk of the evidence base for ODX resides\textsuperscript{1,4,5} it is unclear to what extent, if any, the less commonly-presenting LN+ population may benefit from ODX testing in the adjuvant treatment setting.

RESEARCH QUESTIONS

1. In the adjuvant treatment setting, what is the clinical effectiveness of Oncotype DX in women and men with ER-positive, HER2-negative early stage breast cancer who are lymph node-positive?

2. What are the guidelines associated with Oncotype DX in women and men with ER-positive, HER2-negative early stage breast cancer who are lymph node-positive?
KEY FINDINGS

At present, the evidence base for the use of Oncotype DX in the population of early invasive breast cancer that is ER+, HER2-, and LN+ for guiding treatment decisions in the adjuvant setting is very limited and further restricted to women. Guidelines on the use of Oncotype DX in clinical practice were limited to a few general statements about molecular testing.

METHODS

Literature Search Strategy

A limited literature search was conducted on key resources including PubMed, The Cochrane Library (2013, Issue 12), University of York Centre for Reviews and Dissemination (CRD) databases, Canadian and major international health technology agencies, as well as a focused Internet search. No filters were applied to limit the retrieval by study type. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between January 1, 2008 and December 18, 2013.

Selection Criteria and Methods

One reviewer screened the titles and abstracts of the retrieved publications and evaluated the full-text publications for the final article selection, according to selection criteria presented in Table 1.

Table 1: Selection Criteria

| Population                                                                 | Women and men, ER-positive (ER+) and HER2-negative with early stage breast cancer, lymph node-positive (LN+, 1-3 nodes) |
| Index test                                                                | Oncotype DX                                                                                           |
| Comparator                                                                | None                                                                                                   |
| Outcomes                                                                  | Clinical effectiveness                                                                                 |
| Study Designs                                                              | HTA/ Systematic review/Meta-analysis                                                                   |
|                                                                          | Randomized controlled trials                                                                          |
|                                                                          | Non-randomized studies                                                                                 |
|                                                                          | Guidelines                                                                                             |

Exclusion Criteria

Studies were excluded if, in the event of a mixed population (i.e., LN- and LN+), results were not presented separately for LN+ patients. Likewise, to be considered for inclusion into this review, clinical validity or utility had to be specified as outcomes.

Critical Appraisal of Individual Studies

The AMSTAR instrument⁶ was used to critically appraise the methodological quality of the health technology assessment included in this review while the AGREE II instrument⁷ was used to critically appraise the included set of clinical practice guidelines. No randomized controlled trials were identified from the literature review.
An annotated critical appraisal of the strengths and limitations of the individual included studies is provided in Appendix 3.

SUMMARY OF EVIDENCE

Quantity of Research Available

The literature search yielded 222 citations. After screening titles and abstracts, 196 articles were excluded and 26 potentially relevant articles were selected for full-text review. Five relevant citations were identified from the grey literature. Of these 31 reports, 29 did not meet the inclusion criteria and were excluded, leaving a total of two relevant reports, one of which was a health technology assessment and the other, a set of clinical practice guidelines. No randomized controlled trials were identified that met the inclusion criteria for this review. The study selection process is outlined in Appendix 1. On-going clinical trials of relevance to this topic area are provided in Appendix 5.

Summary of Study Characteristics

Characteristics of the included studies are summarized below and detailed in Appendix 2.

Country of origin

A single health technology assessment by Ward et al.1 of the UK-based National Institute for Health Research (NIHR) was identified for this review. Within this HTA, three studies9-11 from the UK, North America, and the US, respectively, specifically examined the LN+ population in isolation.

A single set of clinical practice guidelines issued by the European Society for Medical Oncology (ESMO), which included limited guidance on the use of gene expression profile tests such as Oncotype DX, was identified.

Population

The health technology assessment (HTA) by Ward et al.1 included all patients with early invasive breast cancer in the adjuvant treatment setting. The majority of breast cancers were ER+, LN-. The mean age of the overall patient population ranged between 50 to 60 years. Within this HTA, three studies9-11 specifically examined the LN+ population in isolation: Goldstein et al.11 (n=465, LN-/LN+; 1-3 nodes: 43.6%), Dowsett et al.9 (n=1231, LN-/LN+; LN+: 25%), and Albain et al.10 (n=367, all N+; 1-3 nodes: 61.9%); none of the trials studied men.

The ESMO guidelines updated the diagnosis, treatment, and follow-up of primary breast cancer.

Index test

Ward et al.1 reviewed nine gene expression profiling (GEP) or immunohistochemistry (IHC) tests – including Oncotype DX, a GEP test – used to support decision-making about treatments in early breast cancer. No reference standard was identified for any of the tests under review, including Oncotype DX.
Information about molecular tests was very limited in the ESMO guidelines;\(^8\) Oncotype DX, nonetheless, was cited within a few general statements made about testing.

**Comparators**

Comparators specifically targeted in the Ward et al.\(^1\) review were those used in current clinical practice in the UK, such as Adjuvant! Online and/or the Nottingham Prognostic Index; however, other comparators were eligible for inclusion.

There was no specific review of the various molecular tests, nor of any potential comparators in the ESMO guidelines.\(^8\)

**Outcomes**

Analytical validity, clinical validity, and clinical utility comprised the included effectiveness outcomes in Ward et al.,\(^1\) of which clinical validity and clinical utility were specifically identified as relevant to this Rapid Response report. Definitions of each outcome can be found in the Glossary of Terms in Appendix 6.

In the ESMO guidelines,\(^8\) desired outcomes from molecular testing related to the possibility of enhanced prognostic and predictive information to support decision-making around treatments.

**Summary of Critical Appraisal**

The health technology assessment by Ward et al.\(^1\) was generally complied with the AMSTAR\(^6\) checklist. The design was adequately described, the literature search was comprehensive, the included and excluded studies were presented clearly, and the quality of included studies was evaluated using validated tools. Study selection and data extraction were performed by a single reviewer, however, with any uncertainty or discrepancies resolved with the assistance of a second or third reviewer. Although the comparator was current UK clinical practice (i.e., Adjuvant! Online or the Nottingham Prognostic Index along with usual clinicopathological indicators), other comparators were potentially eligible. A risk of publication bias was identified owing to the number of sponsored studies. No reference standard was identified for any of the nine tests under study, including Oncotype DX. No meta-analysis was performed because of significant heterogeneity between the included studies, which prevented pooling of the data; thus, findings were presented in narrative format only.

The clinical practice guidelines from the European Society for Medical Oncology (ESMO)\(^8\) included in this review were an update on the diagnosis, treatment, and follow-up of primary breast cancer. It was challenging to assess the guideline development process (e.g., scope and purpose of the guideline, comprehensiveness of the literature search, scientific rigor, stakeholder engagement, funding source[s]) as this was not described in detail. The approach for grading the evidence behind the recommendations was reported as being concordant with that of the American Society of Clinical Oncology (ASCO); however, no supplementary information outlining this approach was included. Guidance around the use of molecular testing was limited to a few general statements without reference to any specific test, though Oncotype DX was included in the examples of molecular tests cited.
Summary of Findings

In the adjuvant treatment setting, what is the clinical effectiveness of Oncotype DX in women and men with ER-positive, HER2-negative early stage breast cancer who are lymph node-positive?

For this review, clinical effectiveness was considered a composite outcome consisting of clinical validity and clinical utility. A single UK National Institute for Health Research (NIHR) health technology assessment (HTA) by Ward et al.\(^1\) published in October 2013, which reviewed nine gene expression profiling and expanded immunohistochemistry tests used in the adjuvant treatment setting of breast cancer, met the inclusion criteria for this review. In addition to identifying new evidence, Ward et al.\(^1\) summarized two previous systematic reviews on the topic by Marchionni et al. (all LN- studies)\(^4\) and Smartt. (mix of LN- and LN+ studies)\(^5\). A total of three trials\(^9-11\) from the HTA were identified that looked at the LN+ population in isolation; the remainder of the evidence applied to the LN- or the undifferentiated (LN-/LN+) population. Smartt\(^5\) included a nested case control study by Goldstein et al.,\(^11\) (n=465, LN-/LN+; 1-3 nodes: 43.6%) which examined clinical validity; in the subgroup of LN+ patients, ODX was found to better predict relapse at 5 years in chemotherapy/hormonal therapy-treated patients than usual clinical features. Since the review by Marchionni et al.\(^4\) consisted entirely of LN- studies, none of its constituent studies were examined in this review. Two other retrospective cohort studies\(^9,10\) were identified by Ward et al.\(^1\). Dowsett et al.\(^9\) (n=1231, LN-/LN+; LN+: 25%) also examined clinical validity in the subgroup of LN+ patients and found that the ODX recurrence score (RS, defined in Appendix 6) was significantly associated with time to distant recurrence (HR 3.47, 95% CI 1.64 to 7.38; \(P < 0.002\)). Albain et al.\(^10\) looked at clinical utility in an exclusively LN+ population. RS was found to be prognostic in the tamoxifen alone group (HR 2.64, 95% CI 1.33 to 5.27; \(p = 0.006\)); there was no benefit of chemotherapy found with a low RS, but improved disease-free survival when RS was high (adjusted HR 0.59, 95% CI 0.35 to 1.01; \(P = 0.033\)).

Only one set of guidelines from the European Society for Medical Oncology (ESMO)\(^8\) was identified from the literature review. Guidance on molecular testing was largely uninformative and limited to a few general statements; Oncotype DX was, however, included among the examples of molecular tests cited. As predictors of prognosis, molecular tests were characterized as still being under investigation and thus, their appropriate integration into existing prognostic models remained to be elucidated.

What are the guidelines associated with Oncotype DX in women and men with ER-positive, HER2-negative early stage breast cancer who are lymph node-positive?

One set of guidelines from the European Society for Medical Oncology (ESMO)\(^8\) was identified from the literature review. No specific guidance was provided around the use of Oncotype DX; rather, only a few general statements were made about molecular testing. Oncotype DX was, however, included in the examples of molecular tests cited within these general statements. This guideline did not include any specific guidance for men with breast cancer.

Limitations

The evidence base for the use of Oncotype DX resides principally within the LN- population; studies focusing exclusively on the LN+ population are scarce. Likewise, men are an underrepresented group within Oncotype DX studies of either the LN- or LN+ population. The
use of molecular testing in breast cancer, however, is an active area of research, and the topic will likely require frequent updating as the evidence base grows.

CONCLUSIONS AND IMPLICATIONS FOR DECISION OR POLICY MAKING

The clinical effectiveness of Oncotype DX, as defined by its clinical validity and clinical utility in the population of early invasive breast cancer that is ER+, HER2-, and LN+, remains uncertain as three trials were identified and limited by their retrospective designs. Lymph node-positive status was defined as having between one to three nodes; however, the majority of the evidence base resided within the lymph node-negative population. Further, the evidence base for LN+ was restricted to women with breast cancer. Finally, guidelines on molecular testing were scarce and largely uninformative. The scientific methodology of the included HTA from which the three trials were derived was sufficient in quality while that of the included guideline was difficult to assess owing to a lack of reporting. Since molecular testing in breast cancer is an active area of research, it is expected that this evidence base will continue to grow; accordingly, more informative data may be forthcoming.

PREPARED BY:
Canadian Agency for Drugs and Technologies in Health
Tel: 1-866-898-8439
www.cadth.ca
REFERENCES


APPENDIX 1: Selection of Included Studies

222 citations identified from electronic literature search and screened

196 citations excluded

26 potentially relevant articles retrieved for scrutiny (full text, if available)

5 potentially relevant reports retrieved from other sources (grey literature, hand search)

31 potentially relevant reports

29 reports excluded:
- irrelevant population (6)
- irrelevant outcomes (10)
- already included in at least one of the selected systematic reviews (7)
- other (review articles, editorials, summaries) (6)

2 reports included in review
### Appendix 2: Summary of Study Characteristics

<table>
<thead>
<tr>
<th>First Author, Publication Year, Country</th>
<th>Study Design</th>
<th>Population</th>
<th>Index test</th>
<th>Comparator</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ward, 2013, UK</td>
<td>HTA</td>
<td>Early, invasive breast cancer in adjuvant treatment setting (majority: ER+ LN-); mean age 50-60 years.</td>
<td>Nine GEP or expanded IHC tests, including ODX</td>
<td>Current UK clinical practice (i.e., Adjuvant! Online and/or Nottingham Prognostic Index; other comparators, however, were eligible for inclusion)</td>
<td>Clinical validity; clinical utility</td>
</tr>
</tbody>
</table>

Three included studies within HTA examined LN+ in isolation: Goldstein et al.\(^{11}\) (n=465, LN-/LN+; 1-3 nodes: 43.6%), Dowsett et al.\(^{9}\) (n=1231, LN-/LN+; LN+: 25%), and Albain et al.\(^{10}\) (n=367, all N+; 1-3 nodes: 61.9%); no men were studied; only Albain et al.\(^{10}\) included Canadian patients.

---

ER=Estrogen receptor; GEP=Gene expression profiling; HTA=health technology assessment; IHC=immunohistochemistry; LN=Lymph node; ODX=Oncotype DX; UK=United Kingdom
## Appendix 3: Summary of Critical Appraisal

<table>
<thead>
<tr>
<th>First Author, Publication Year, Country</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HTA/Systematic review/Meta-analysis</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Ward, 2013, UK | • Design was *a priori* with comparators pre-specified  
• Comprehensive literature search performed  
• A listing of included and excluded studies was provided; characteristics of included studies were described while reasons for exclusion were given for excluded studies.  
• Although controlled studies were targeted, the best level of available evidence was included; a quality assessment was applied to all included studies and  
• Results were presented in narrative form as degree of heterogeneity prevented meta-analysis  
• Scientific quality was used in formulating conclusions about the included studies  
• Declarations of interest and sources of financial support both reported | • Study selection and data extraction performed by single reviewer (uncertainty or discrepancies resolved with second or third reviewer)  
• Literature search limited to English language articles (unless no other comparable publications)  
• No reference standard for ODX  
• Comparator was ‘current UK clinical practice’  
• LN+ defined as having up to 3 nodes  
• Three studies reported on LN+ population; majority of evidence base is for LN- population  
• A risk of publication bias and possible conflict of interest were cited because of a number of sponsored studies; among the three LN+ studies, all received funding from the ODX manufacturer. |

### Clinical practice guidelines

| Aebi, 2011, Europe (ESMO) | • Although not explicitly stated at outset, guideline appears to target a clinical audience  
• Guideline appears to have been developed by experts from the clinical oncology community; competing interests are described for lead authors  
• A note was provided at the end of the guideline indicating that levels of evidence [I-V] and grades of recommendation [A-D] were adopted according to the approach of the ASCO; recommendations based on expert opinion were provided without brackets.  
• Health benefits, side effects, and risks appear to have been taken into account in formulating the | • Scope and purpose of guideline not described at outset: no description of objective of guideline or health question(s) to be covered; target population likewise not characterized. A search through guideline revealed that men as a patient population were not addressed.  
• It is unclear to what extent, if any, stakeholders outside of the clinical oncology community (e.g., patient groups) were engaged in the development of the guideline  
• The scientific rigor underlying the guideline development process was not described: no information was provided on the literature search methodology, the criteria used for evidence selection, or the quality of the included evidence.  
• It is unclear whether the guideline was |
<table>
<thead>
<tr>
<th>First Author, Publication Year, Country</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
</table>
|                                        | recommendations; however, there is often no explicit link made to supporting evidence.  
- Recommendations are explicit; however, they are not so easily identifiable from the body of text in which they are embedded. Different or alternative management options are likewise clearly described in the text. | externally reviewed prior to publication; likewise, the process for updating the guideline was not described, though a note was included to indicate that the present guideline superseded a previous version from 2010.  
- There was no information provided on the source of funding for the guideline. |

ASCO= American Society of Clinical Oncology; ESMO= European Society for Medical Oncology; HTA= health technology assessment; LN= Lymph node; ODX= Oncotype DX; UK= United Kingdom; US= United States
## Appendix 4: Summary of Findings

<table>
<thead>
<tr>
<th>First Author, Publication Year, Country, Study Design</th>
<th>Main Study Findings</th>
<th>Authors’ conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HTA/Systematic review/Meta-analysis</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Ward, 2013, UK, HTA                                 | Two additional studies\(^5,10\) reported on LN+ population (one on clinical validity,\(^9\) the other on clinical utility\(^10\)); one other study\(^11\) looking at clinical validity was reported in a previous systematic review\(^5\) (included in this HTA)  
- Ward et al. included two earlier systematic reviews by Marchionni et al.\(^4\) and Smartt\(^5\); the former systematic review consisted only of LN- studies while the latter consisted of a single LN+ study.\(^11\)  
  - In Smartt's Clinical validity: One study by Goldstein et al.,\(^11\) of LN+ patients found ODX to better predict relapse at 5 years in CMT/HT-treated patients than usual clinical features; there were no data for clinical utility.  
- Ward et al. also identified two additional LN+ studies,\(^9,10\) updating the overall ODX evidence base from the two previous systematic reviews\(^4,5\):  
  - Clinical validity: Dowsett et al.\(^9\) found ODX RS significantly associated with time to distant recurrence (HR 3.47, 95% CI 1.64 to 7.38; \(P < 0.002\)); there were no data for clinical utility.  
  - Clinical utility: Albain et al.\(^10\) found RS prognostic in tamoxifen alone group (HR 2.64, 95% CI 1.33 to 5.27; \(p = 0.006\)); no benefit of CAF with low RS, but improved DFS in high RS (HR\(^*\) 0.59, 95% CI 0.35 to 1.01; \(P = 0.033\)); there were no data for clinical validation. | From Ward et al.'s review of the two previous systematic reviews\(^4,5\):  
- "Clinical validity (prognostic ability of the tests) – there is fairly strong support for Oncotype DX over and above standard clinical predictors, but only in a well-defined population (ER+, LN-)." (p.24)  
- "Clinical utility – very few of the studies, particularly in isolation, provided compelling evidence of the test’s clinical utility.” (p.24)  
From the two additional LN+ studies\(^9,10\) in Ward et al.:  
- Dowsett et al.\(^9\): Clinical validity: “… the findings demonstrated that RS is an independent predictor of distant recurrence in LN– and LN+ hormone receptor-positive patients treated with anastrozole, adding value to estimates with standard clinicopathological features.” (p. 32)  
- Albain et al.\(^10\): Clinical utility: “…RS is prognostic for tamoxifen-treated patients with positive nodes and predicts significant benefit of CAF in tumours with a high RS. A low score identifies women who might not benefit from anthracycline-based chemotherapy, despite positive nodes.” (p. 36) |
| **Clinical practice guideline**                      |                     |                      |
| Aebi, 2011, Europe (ESMO)                           | Specific guidance on the use of gene expression profile tests was limited.  
- Gene expression profiles (e.g., ODX) may provide extra prognostic and/or predictive data to enhance insights from pathology findings and possibly identify patients most likely to benefit from specific adjuvant treatment strategies. | “Molecular predictors of prognosis (e.g., ODX) may outperform the traditional prognostic markers in certain patient populations, but their integration into prognostic models is not yet established.” (p. 13) |
<table>
<thead>
<tr>
<th>First Author, Publication Year, Country, Study Design</th>
<th>Main Study Findings</th>
<th>Authors’ conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>likely to benefit from adjuvant chemotherapy, particularly from the ER-positive, early breast cancer population.</td>
<td>models is still under investigation.” (p.vi21)</td>
</tr>
</tbody>
</table>

CAF=cyclophosphamide/doxorubicin/5-fluorouracil; CMT=chemotherapy; DFS=disease-free survival; ER=estrogen receptor; ESMO=European Society for Medical Oncology; HT=hormonal therapy; HR=hazard ratio; HTA=health technology assessment; LN=lymph node; ODX=Oncotype DX; RS=recurrence score; UK=United Kingdom; US=United States

*adjusted for number of positive nodes
Appendix 5: List of On-going Clinical Trials


WHO International Clinical Trials Registry Platform [Internet]. Geneva: World Health Organization. Identifier EUCTR2012-000576-42-IE, S1007 is a clinical trial for patients with node positive (1-3 positive nodes) Hormone Receptor-Positive and HER2-Negative breast cancer who when tested with the oncotype DX test are shown to have a recurrence score (RS) of 25 or less. Patients will be randomised to receive endocrinetherapy +/- chemotherapy; 2013 Jul 15 [cited 2014 Jan 22]. Available from: http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2012-000576-42-IE

Appendix 6: Glossary of Terms
Unless otherwise indicated, the following terms were defined according to Ward et al.:¹

**Analytical validity:** “the ability of the test to accurately and reliably measure the expression of mRNA or proteins by breast cancer tumor cells, that is, repeatability and reproducibility.” (p.17)

**Clinical validity:** “the degree to which the test can accurately predict the risk of an outcome (typically distant recurrence) and discriminate patients with different outcomes. This relates to the prognostic ability of the test – does the test have evidence on clinical validity and has this been externally validated (in an independent data set).” (p.17)

**Clinical utility:** “the ability of the test to discriminate between those who will have more or less benefit from a therapeutic intervention. [This] relates to improvements in clinical outcomes such as overall survival, disease-free survival, chemotherapy toxicity or quality of life.” (p.17)

**Recurrence score²:** “a number between 0 and 100 that corresponds to a specific likelihood of breast cancer recurrence within 10 years of the initial diagnosis, as well as response to adjuvant treatment.” (p.1)