Extended executive summary

Executive Summary, extended methods and results

BACKGROUND

Breast cancer is the second most common cancer in Switzerland, with approximately 6,300 women and 50 men being diagnosed annually. The disease arises from genetic mutations leading to uncontrolled cell growth in breast tissue. Diagnosis involves physical examination, imaging, and histopathological analysis, with staging based on tumour characteristics and biomarker expression. Treatment includes surgery, radiotherapy, endocrine therapy, and chemotherapy, guided by prognostic factors and patient-specific considerations. While earlier stages have better prognosis, recurrence remains a risk. Adjuvant chemotherapy (chemotherapy following surgery) reduces recurrence risk but can lead to adverse events, necessitating a decision-making approach based on individual patient profiles. Multigene-expression tests, which assess genes related to proliferation and the oestrogen receptor pathway, have been developed and are primarily used to assist with recurrence risk prediction and, in some cases, to evaluate the potential benefit of adjuvant chemotherapy. In Switzerland, multigene-expression tests are temporarily covered by the mandatory health insurance since 2015 for patients with ER+/HER2- breast cancer with up to 3 affected lymph nodes, for whom the results of conventional testing alone do not allow a clear decision to be made regarding adjuvant chemotherapy. The 4 multigene-expression tests that are currently covered are: Oncotype DX, MammaPrint, EndoPredict, and Prosigna.

OBJECTIVE

This health technology assessment (HTA) report assesses the efficacy, effectiveness, safety, cost-effectiveness and budget impact as well as ethical, legal, social, and organisational benefits and harms of multigene-expression tests when applied as described in the reimbursement texts in Switzerland.

1. Clinical effectiveness

METHODS

A systematic literature search was conducted in PubMed (MEDLINE), Embase.com and Cochrane Library on 6 May 2024. Studies were selected by applying pre-specified inclusion criteria, which were based on the PICO, study design and type of evidence (i.e. comparative clinical effectiveness[1] and the intermediate measures predictive ability[a], prognostic ability[a],

[1] See *Fehler! Verweisquelle konnte nicht gefunden werden.* ***main report***

impact on treatment management[a]). Outcomes[a] included survival outcomes (e.g. overall survival), health-related quality of life (HRQoL) and change in treatment management. A stepwise selection approach was implemented for studies reporting prognostic ability: (1) selection of prospective RCTs; (2) if no prospective RCT was found, other study designs reporting prognostic ability were selected. Prospective RCTs, retrospective/re-analyses of RCTs and comparative non-randomised studies were selected for comparative clinical effectiveness and predictive ability. The risk of bias of includes studies was critically appraised with RoB 2, adapted PROBAST or the adapted ROBINS-I tool. The overall certainty of the evidence was assessed with GRADE. The included studies were analysed separately for each multigene-expression test. For the outcome impact on treatment management pooled event rates were calculated by meta-analysis.

RESULTS

In total, 25 studies on Oncotype DX (including 2 RCTs: RxPONDER and TAILORx), 9 studies on MammaPrint (including 1 RCT: MINDACT), 7 studies on EndoPredict, and 4 studies on Prosigna were included in the systematic review. An ongoing RCT was identified for Prosigna (OPTIMA). Of the 4 multigene-expression tests, only for MammaPrint evidence is available on the comparative clinical effectiveness of the multigene-expression test versus conventional testing for guiding adjuvant chemotherapy decisions in early breast cancer patients. The other studies on MammaPrint and the studies on Oncotype DX, EndoPredict and Prosigna reported intermediate measures on the predictive ability, prognostic ability, and impact on treatment management. The intermediate measures predictive and prognostic ability lack the comparison between the multigene-expression test versus conventional testing and in the impact on treatment management follow-up of clinical outcomes is not included. The predictive and prognostic ability of the multigene-expression tests were based on different survival outcomes. No data was reported on HRQoL. Except for a small number of studies on the impact on treatment management, none of the studies included the population for whom it was unclear based on conventional testing whether to prescribe adjuvant chemotherapy.

Oncotype DX

Comparative clinical effectiveness: No studies or ongoing RCTs identified.

Intermediate measure – predictive ability/chemotherapy benefit: Nine included studies reported on the predictive ability of Oncotype DX for chemotherapy benefit on different survival outcomes.

Early breast cancer patients classified by Oncotype DX in genomic low or intermediate risk may have little to no chemotherapy benefit on overall, breast cancer-specific, disease-free, recurrence-free, invasive disease-free, and distant metastasis-free survival (2 RCTs: moderate to low certainty evidence & 7 studies: very low certainty evidence). These two prospective RCTs provided the best available evidence in patients classified as Oncotype DX intermediate and low/intermediate risk. In the TAILORx trial no statistically significant chemotherapy effect was

observed on 9 years overall survival (adjusted hazard ratio [aHR] 1.01; 95% confidence interval [CI] 0.82-1.27), invasive disease-free survival (aHR 0.93; 95% CI 0.81-1.06) and distant metastasis-free survival (aHR 0.91; 95% CI 0.71-1.18), and a statistically non-significant chemotherapy benefit was observed on recurrence-free survival (aHR 0.90; 95% CI 0.73-1.11), in women with HR+, HER2-, zero involved lymph nodes early breast cancer and an intermediate recurrence score of 11-25 (1 RCT: moderate certainty evidence). The absolute differences in survival rates between women who received chemo-endocrine therapy versus endocrine therapy ranged from 0.1% to 1.0%. In the RxPONDER trial a statistically non-significant chemotherapy benefit was observed on 5 years invasive disease-free survival (aHR 0.86; 95% CI 0.72-1.03) and distant metastasis-free survival (aHR 0.88; 95% CI 0.71-1.09) in women with HR+, HER2-, 1-3 involved lymph nodes early breast cancer and a low/intermediate recurrence score of 0-25 (1 RCT: low certainty evidence). The absolute differences in survival rates for chemo-endocrine therapy versus endocrine therapy were 1.2% for invasive disease-free survival and 1.0% for distant metastasis-free survival. Results of 7 other studies in early breast cancer patients classified as Oncotype DX low to intermediate risk were mostly in line with the 2 prospective RCTs (7 studies: very low certainty evidence).

Three of these 7 studies included patients classified as Oncotype DX high risk. Early breast cancer patients classified by Oncotype DX in genomic high risk may have a chemotherapy benefit on overall, disease-free, recurrence-free, invasive disease-free, and distant metastasis-free survival, but the evidence is very uncertain (3 studies: very low certainty evidence).

Intermediate measure – prognostic ability: Oncotype DX may be prognostic for invasive disease-free survival in early breast cancer patients (1 RCT: low certainty evidence). The RxPONDER trial showed that a higher recurrence score was related to a statistically significant increased risk of invasive disease events (aHR 1.05; 95% CI 1.04-1.07) in early breast cancer patients classified as low/intermediate risk with a recurrence score of 0-25.

Intermediate measure – impact on treatment management: In women and men with ER+ or HR+, HER2-, 0-3 involved lymph nodes early breast cancer, Oncotype DX may change the adjuvant treatment recommendation in 33.5% of the patients, but the evidence is very uncertain (95% CI 27.6-40.1%; 12 studies: very low certainty evidence): from endocrine to chemotherapy in 7.0% (95% CI 5.4-9.0%) and from chemo to endocrine therapy in 25.5% (95% CI 19.4-32.8%). In women with clinically intermediate risk ER+, HER2-, 0-3 involved lymph nodes early breast cancer, these percentages were respectively 23.2%, 3.9% and 18.9%, but the evidence is very uncertain (4 studies: very low certainty evidence).

MammaPrint

Comparative clinical effectiveness: The design of the MINDACT trial is aimed at assessing the clinical outcomes of treating early breast cancer patients with adjuvant chemotherapy based on their clinical risk (by Adjuvant!Online) or their genomic risk (by MammaPrint). Patients with

discordant risks (i.e. C-high/G-low or C-low/G-high) were randomised to chemotherapy or no chemotherapy. In a prespecified secondary interim analysis not based on these discordant risk groups, the comparative clinical effectiveness of MammaPrint versus Adjuvant!Online was estimated by comparing distant metastasis-free survival resulting from treating according to genomic risk strategy or clinical risk strategy. Treatment based on MammaPrint may result in 46% fewer clinical high risk early breast cancer patients being treated with adjuvant chemotherapy and little difference in distant metastasis-free survival (1 RCT: low certainty evidence). The interim 8 years distant metastasis-free survival of adjuvant treatment based on MammaPrint compared to the clinical risk strategy was 90.9% versus 91.3%, and this slightly shorter survival becomes larger at 10 years follow-up. Whether adjuvant treatment based on MammaPrint is clinically beneficial depends on weighing the benefits (i.e. less adverse events due to adjuvant chemotherapy) and harms (i.e. possibly a slightly shorter distant metastasis-free survival) in a contextualised appraisal. This is not reported in the current MINDACT publications.

Intermediate measure – predictive ability/chemotherapy benefit: Three studies reported on the predictive ability of MammaPrint for chemotherapy benefit on different survival outcomes.

Early breast cancer patients classified by Adjuvant!Online and MammaPrint in C-low/G-high may have little to no chemotherapy benefit on overall, disease-free and distant metastasis-free survival (1 RCT: low certainty evidence). In the MINDACT trial, classification in C-low/G-high may not result in a statistically significant chemotherapy effect on 8 years overall survival (0.94; 95% CI 0.54-1.67) and may result in a statistically non-significant chemotherapy benefit on 8 years disease-free survival (aHR 0.79; 95% CI 0.55-1.13) and distant metastasis-free survival (aHR 0.85; 95% CI 0.53-1.37). Early breast cancer patients classified by Adjuvant!Online and MammaPrint in C-high/G-low may have a chemotherapy benefit on distant metastasis-free survival and may have little to no chemotherapy benefit on overall and disease-free survival (1 RCT low certainty evidence). Classification in C-high/G-low may result in a statistically significant chemotherapy benefit on distant metastasis-free survival (aHR 0.66; 95% CI 0.48-0.92), and a statistically non-significant chemotherapy benefit on 8 years overall survival (aHR 0.69; 95% CI 0.45-1.05) and disease-free survival (aHR 0.79; 95% CI 0.62-1.02).

MammaPrint may not be predictive for chemotherapy benefit on breast cancer-specific survival in early breast cancer patients, but the evidence is very uncertain (1 study: very low certainty evidence). A pooled analysis of 6 patient series in women with stage T1-3, 0-3 involved lymph nodes early breast cancer did not show statistically significant (p=0.45) differences in chemotherapy benefit on 5 years breast cancer-specific survival in women classified as MammaPrint low risk (aHR ∞; 95% CI 0-∞) and high risk (aHR 0.21; 95% CI 0.06-0.80).

A database study reported inconsistent and very uncertain results with a statistically non-significant benefit of no chemotherapy on 5 years overall survival in women with ER+, HER2-

invasive lobular carcinoma breast cancer classified as MammaPrint high risk (aHR 1.41; 95% CI 0.44-4.55; 1 study: very low certainty evidence).

Intermediate measure – prognostic ability: MINDACT showed that MammaPrint may be prognostic for overall, disease-free and distant metastasis-free survival in early breast cancer patients, but the evidence is very uncertain (1 RCT: very low certainty evidence). In women with the same clinical risk and adjuvant treatment, survival followed the genomic risk classification in the discordant risk groups. In women who did not receive chemotherapy, survival was lower in C-low/G-high patients (respectively 93.0%; 81.9%; 90.8%) compared with C-low/G-low patients (respectively 96.5%; 86.8%; 94.7%). In women who received chemotherapy, survival was higher in C-high/G-low patients (respectively 95.7%; 86.4%; 92.0%) compared with C-high/G-high patients (respectively 90.1%; 79.1%; 85.9%).

Intermediate measure – impact on treatment management: In women with ER+ or HR+, HER2-, 0-3 involved lymph nodes early breast cancer, MammaPrint may change the adjuvant treatment recommendation in 34.9% of the patients, but the evidence is very uncertain (95% CI 26.4-44.5%; 5 studies: very low certainty evidence): from endocrine to chemotherapy in 13.3% (95% CI 11.5-15.3%) and from chemo to endocrine therapy in 20.7% (95% CI 12.5-32.2%). In women with HR+, HER2-, 0-3 involved lymph nodes early breast cancer and an intermediate Oncotype DX result, these percentages were respectively 33.6%, 20.5% and 13.1%, but the evidence is very uncertain (1 study: very low certainty evidence).

EndoPredict

Comparative clinical effectiveness: No studies or ongoing RCTs identified.

Intermediate measure – predictive ability/chemotherapy benefit: In women with ER+, HER2-, up to 10+ involved lymph nodes early breast cancer EndoPredict may be predictive for chemotherapy benefit on 10 years breast cancer recurrence and distant recurrence, but the evidence is very uncertain (1 study: very low certainty evidence). The increase in breast cancer recurrence risk with EPclin score was statistically significant reduced (p-interaction=0.025) in women who received chemo-endocrine therapy (aHR 2.06; 95% CI 1.82-2.34) versus endocrine therapy (aHR 2.50; 95% CI 2.26-2.76). Similar results were reported for distant recurrence.

Intermediate measure – prognostic ability: EndoPredict may be prognostic for breast cancer and distant recurrence in early breast cancer patients, but the evidence is very uncertain (3 studies: very low certainty evidence).

The adjusted hazard ratios for 10 years breast cancer and distant recurrence reported above can be interpreted also as prognostic ability: a higher Epclin score was related to a statistically significant increased recurrence risk in women who received chemo-endocrine therapy as well as in women who received endocrine therapy (1 study; low certainty evidence).

In women with ER+, HER2-, up to 4+ involved lymph nodes early breast cancer treated with endocrine therapy, the 10 years distant recurrence rate followed the genomic risk classification in the discordant clinical-genomic risk groups and EndoPredict may be prognostic for 10 years distant recurrence survival, but the evidence is very uncertain (2 studies; very low certainty evidence).

Intermediate measure – impact on treatment management: In women with clinically intermediate risk ER+, HER2-, 0-3 involved lymph nodes early breast cancer, EndoPredict may change the adjuvant treatment recommendation in 39.6% of the patients, but the evidence is very uncertain (95% CI 35.0-44.3%; 4 studies: very low certainty evidence): from endocrine to chemotherapy in 8.6% (95% CI 3.9-17.8%) and from chemo to endocrine therapy in 28.6% (95% CI 18.8-41.0%).

Prosigna

Comparative clinical effectiveness: No studies identified. An RCT is ongoing with an estimated completion date in December 2034. Five-year follow-up data is anticipated to be published mid-2026.

Intermediate measure – predictive ability/chemotherapy benefit: No studies identified.

Intermediate measure – prognostic ability: In women with HR+, HER2-, pT1pN0 tumours who were not recommended any adjuvant treatment, Prosigna may be prognostic for 15 years breast cancer-specific survival and 10 years distant metastasis-free survival, but the evidence is very uncertain (1 study: very low certainty evidence). A statistically significant increased risk of breast cancer-related death was observed in women classified as G-intermediate versus G-low (aHR 4.52; 95% CI 1.08-18.85) and in G-high versus G-low (aHR 9.09; 95% CI 1.80-14.50). Comparable, but more imprecise results were reported for distant metastasis-free survival.

Intermediate measure – impact on treatment management: In post-menopausal women with ER+, HER2- and zero involved lymph nodes early breast cancer, Prosigna may change the adjuvant treatment recommendation in 17.1% of the patients, but the evidence is very uncertain (95% CI 13.8-21.1%; 3 studies: very low certainty evidence): from endocrine to chemotherapy in 11.1% (95% CI 8.8-13.9%) and from chemo to endocrine therapy in 5.4% (95% CI 2.2-12.5%).

2. Costs, cost-effectiveness and budget impact

METHODS

For the economic review, a systematic literature search was conducted in PubMed (MEDLINE), Embase.com, Cochrane Library, Tufts Medical Centre Cost-Effectiveness Analysis (CEA) Registry, and National Health Service Economic Evaluation Database (NHS EED). Studies were selected by applying pre-specified inclusion criteria based on the PICO and study design. The quality of the studies was assessed using the Philips checklist for health economic models.

Based on this quality assessment, one health economic model was selected to be replicated and adapted to the Swiss setting by incorporating Swiss data on population characteristics, survival, and treatment costs. An economic analysis taking a healthcare payers' perspective and lifetime horizon was conducted in a target population representing the Swiss HR+, HER2-, LN0-3 early breast cancer population, as well as an intermediate risk population which was used to represent the population as defined in the reimbursement text (with the exception of MammaPrint, for which this analysis was not possible due to data constraints).

RESULTS

In the economic review 36 studies on multigene-expression tests for patients with HR+, HER2-, LN0-3 early breast cancer were included: 29 evaluated Oncotype DX, 10 evaluated MammaPrint, 6 evaluated EndoPredict, and 5 evaluated Prosigna. In cost-effectiveness results of the studies, Oncotype DX and MammaPrint ranged from dominant to dominated, while EndoPredict and Prosigna ranged from dominant to positive ICERs.

The results of the economic model assessment indicated that adding a multigene-expression test (i.e. Oncotype DX, MammaPrint, EndoPredict, or Prosigna) to conventional testing is likely to result in very small changes in health outcomes (life years and QALYs). However, the direction and magnitude of the difference in health effects is very uncertain (Oncotype DX: 0.00 QALY [95% CI: -0.07, 0.06], MammaPrint: 0.00 QALY [95% CI: -0.10, 0.11], EndoPredict: 0.03 QALY [95% CI: -0.05, 0.14], Prosigna 0.03 QALY [95% CI: -0.04, 0.12]). This means that based on the available evidence, it cannot be determined whether adding one of the evaluated multigene-expression tests to conventional testing will result in a health gain or loss. This also holds for the incremental health effects of multigene-expression tests in the intermediate risk population (Oncotype DX: -0.01 QALY [95% CI: -0.15, 0.14], EndoPredict: 0.04 QALY [95% CI: -0.13, 0.25], Prosigna: 0.04 QALY [95% CI: -0.11, 0.23]). Oncotype DX, EndoPredict, and Prosigna are expected to result in an increase in healthcare costs due to the additional costs of testing which are not offset by the savings in costs for adjuvant chemotherapy (Oncotype DX: CHF 2'008 [95% CI: -8'502, 7'093] , EndoPredict: CHF 2'778 [95% CI: -8'210, 11'695] Prosigna: CHF 2'127 [95% CI: -8'287, 7'364]). This also holds when multigene-expression tests are only used in the intermediate risk population (Oncotype DX: CHF 327 [95% CI: -9'621, 9'759], EndoPredict: CHF 4202 [95% CI: -8'612, 16'466], Prosigna: CHF 1'664 [95% CI: -8'233, 11'070]). However, the analysis for MammaPrint resulted in a reduction in healthcare costs (CHF -749 [95% CI: -10'118, 9'202]). The sensitivity analysis indicated that the model outcomes are particularly sensitive to changes the recurrence risk and in the estimates of the probability to receive adjuvant chemotherapy under either testing strategy. Swiss data to inform these parameters is lacking. Oncotype DX had a less than 50% probability of being cost-effective at any willingness-to-pay threshold, whereas EndoPredict and Prosigna has a more than 50% probability of being cost-effective at willingness-to-pay levels above CHF 100'000 per QALY gained. MammaPrint had a 56.1% probability of being cost effective at a willingness-to-forgo level of 0 CHF per QALY lost.

The 5-year budget impact analysis indicated that the use of Oncotype DX, EndoPredict, or Prosigna in addition to conventional testing would require an increase of the healthcare budget (Oncotype DX: CHF 28.9M, EndoPredict: 45.5M, Prosigna: 38.6M), whereas a lower healthcare budget would be required when MammaPrint is used (CHF -13M). The budget impact is considerably smaller if multigene-expression tests are only used in the intermediate risk population (Oncotype DX: CHF 1.0M, EndoPredict: 19.6M, Prosigna: 9.7M).

3. Ethical, legal, social and organisational issues

METHODS

Ethical, legal, social, and organisational (ELSO) issues were searched through the systematic literature searches of the clinical effectiveness and cost-effectiveness in PubMed (MEDLINE), Embase.com, Cochrane Library, Tufts Medical Centre Cost-Effectiveness Analysis (CEA) Registry, and National Health Service Economic Evaluation Database (NHS EED) and targeted non-systematic searches. Findings were described narratively.

RESULTS

Fourteen articles on ELSO domains were included. In the ethical domain, inadequate representation of ethnic minorities as well as the exclusion of men from validation trials were discussed. No legal issues were found from the searches. Identified social issues included emotional stress that multigene-expression testing can cause to patients along with barriers that prevent patients from being screened, such as out-of-pocket expenses and privacy concerns. In the organizational domain, several factors i.e. gradual learning, patient preferences, and adherence to guidelines are discussed, providing insight into how they inform oncologists' decisions to order multigene-expression tests and variously shape testing utilization patterns. In addition, challenges related to integrating multigene-expression tests into clinical practice, positioning general practitioners as solely emotional support providers due to the lack of expertise in these tests, as well as the impact of sample contamination and extraction methods on test accuracy were also addressed.

CONCLUSION

Overall, the comparative evidence on multigene-expression tests is sparse. Of the 4 multigene-expression tests, only for MammaPrint an interim analysis of one RCT is available on the comparative clinical effectiveness of the multigene-expression test versus conventional testing for guiding adjuvant chemotherapy decisions in early breast cancer patients. Treatment based on MammaPrint may result in 46% fewer clinical high risk early breast cancer patients being treated with adjuvant chemotherapy and little difference in or a slightly shorter distant metastasis-free survival compared to treatment based on conventional testing. All other evidence is based on the intermediate measures predictive ability/chemotherapy benefit and prognostic ability,

which lack the comparison between the multigene-expression test and conventional testing, and the intermediate measure impact on treatment management, which lacks follow-up of clinical outcomes.

The evidence base is insufficient to conduct a robust cost-effectiveness analysis of the multigene-expression tests Oncotype DX, MammaPrint, EndoPredict and Prosigna compared to conventional testing in early breast cancer patients. In particular, data on the relation between recurrence risk and the probability to receive adjuvant chemotherapy under conventional testing is lacking. The effect of adding multigene-expression tests to conventional testing on costs and health outcomes are very uncertain, and it cannot be determined whether it will result in a health gain or loss, nor in an increase or decrease in healthcare costs.

Finally, the use of multigene is associated with several ethical, social and organisational issues.