NIPH
Norwegian Institute of Public Health

REPORT

2022

HEALTH TECHNOLOGY ASSESSMENT:

Molecular tests for detection of PIK3CA mutations in men and postmenopausal women with HR+/HER2–, locally advanced or metastatic breast cancer

Publisher	Norwegian Institute of Public Health, Division for Health Services
Title	Molecular tests for detection of PIK3CA mutations in men and postmenopausal women with HR+/HER2–, locally advanced or metastatic breast cancer: A Health Technology Assessment.
Norwegian title	Molekylære tester for deteksjon av PIK3CA mutasjoner blant menn og postmenopausale kvinner med HR+/HER2–, lokalavansert eller metastasert brystkreft: Metodevurdering
Authors	Flodgren GM, *Project leader, Senior researcher, Norwegian Institute of Public Health* Hamidi V, *Health economist, Senior Researcher, Norwegian Instirute of Pulic Health* Meneses EJ, Researcher, *Norwegian Institute of Public Health* Bidonde J, Senior Researcher, *Norwegian Institute of Public Health*
ISBN	978-82-8406-285-3
Project number	(ID2019_070)
Type of report	Health Technology Assessment (Metodevurdering)
No. of pages	57 (78 including appendices)
Client	The National System for Managed Introduction of New Health Technologies in Norway
Subject heading (MeSH)	PIK3CA mutations, HR+, HER2-, advanced or metastatic breast cancer, tests, prediction, biomarker, Sanger sequencing, Next Generation Sequencing, Polymerase Chain Reaction, Liquid chip technologies, sensitivity, specificity, concordance, accuracy, advantages, limitations, costs and cost analysis
Citation	Flodgren GM, Hamidi V, Meneses JE, Bidonde J. Molecular tests for detection of PIK3CA mutations in men and postmenopausal women with HR+/HER2–, locally advanced or metastatic breast cancer: A Health Technology Assessment 2022. Norwegian Insitutute of Public Health, 2022.

Key messages

The Norwegian Institute of Public Health was commissioned to evaluate molecular tests for the identification of somatic PIK3CA mutations in men and postmenopausal women with HR+/HER2–, advanced or metastatic breast cancer (BC). Tumours harbouring PIK3CA mutations constitute up to 36% of cases with HR+/HER2- BC, which suggests that around 48 patients may be considered for treatment with PI3k inhibitors (e.g. alpelisib) in Norway each year. Accurate and reliable detection of PIK3CA mutations is important for correctly identifying patients who may benefit from targeted treatment.

We included three original studies that reported concordance between (i) two PCR assays; (ii) two NGS-panels, and (iii) ddPCR and NGS for the detection of PIK3CA mutations. Experts were contacted for cost information. The results of this HTA show that:

- Un-pooled results (3 studies) provided inadequate evidence for test accuracy according to the EGAPP quality tool, and quality of reporting was poor.
- Single studies, with relatively small sample sizes, reported very good concordance between tests, which all used plasma samples(Cohen's k: 0.80 to 0.86)
- For the detection of PIK3CA in isolation, the costs for testing using PCR is less than NGS-panel testing. However, using PCR assays for the detection of additional relevant mutations, will increase total cost. At present, the capital and infrastructure as well as maintenance costs are higher for NGS than PCR.
- Assuming that about 140 patients with metastatic BC are eligible for testing to detect of PIK3CA mutations in Norway each year, the costs were estimated to be approximately NOK 322,000.
- All tests have advantages and limitations, but due to incomplete information a proper comparison was difficult to make. The choice of a suitable test for the detection of PIK3CA mutations depends on accessibility of testing modalities, economic considerations, sample type and risk of false negatives, and turnaround time.
- Future research should focus on conducting larger cohort studies with well-defined patient populations, that follows the patients from testing (or no testing), through treatment and final outcomes. Further, robust and replicable methods, as well as a reporting standard checklist, should be used for increased clarity.

Title:
Molecular tests for detection of PIK3CA mutations in men and postmenopausal women with with HR+/HER2–, locally advanced or metastatic breast cancer

---

Type of publication:
Health Technology Assessment

Commissioner:
The National System for Managed Introduction of New Health Technologies in Norway

---

Updated search:
22.10.2021

Team:
Gerd.M.Flodgren, senior researcher, and team leader
Vida Hamadi, health economist, senior researcher
Jose E Meneses, researcher
Julia, Bidonde, senior researcher
Gyri Hval, research librarian

Clinical experts:
Tormod K Guren, Oslo University Hospital, Emilius AM Janssen, Stavanger University Hospital

Hovedfunn (norsk)

Tittel:
Molekylære tester for deteksjon av PIK3CA mutasjoner hos menn og postmenopausale kvinner med HR+/HER2–, lokalavansert eller metastatisk brystkreft

---

Publikasjonstype:
Metodevurdering

Oppdragsgiver:
Nasjonalt system for innføring av nye metoder i spesialisthelsetjenesten

Søk oppdatert:
22.10.2021

Team:
Gerd.M.Flodgren, senior researcher, and project lead
Vida Hamadi, health economist, senior researcher
Jose Menses, researcher
Julia Bidonde, senior researcher
Gyri Hval, research librarian

Kliniske eksperter:
Tormod K Guren, Oslo Universitetssykehus, Emilius AM Janssen, Stavanger Universitetssykehus

Folkehelseinstituttet har på oppdrag fra Bestillerforum for nye metoder evaluert molekylære tester for identifisering av somatiske PIK3CA mutasjoner hos menn og postmenopausale kvinner med HR+/ HER2–, lokalavansert eller metastasert brystkreft (BC). Svulster som har PIK3CA mutasjoner utgjør opp til 36% av ER+/HER2-BC-tilfellene, noe som tilsvarer rundt 48 pasienter per år i Norge som kan kvalifisere for behandling med Pi3k hemmer (f.eks. alpelisib). Tester som korrekt identifiserer hvilke pasienter som kan ha nytte av målrettet medikamentell behandling er viktige for adekvate behandlingsbeslutninger.

Vi inkluderte tre enkeltstudier som rapporterte konkordans mellom (i) to PCR-assays; (ii) to NGS-paneler, og (iii) ddPCR og NGS for deteksjon av PIK3CA-mutasjoner. Eksperter ble kontaktet for informasjon om ressursbruk. Resultatene av metodevurderingen viser at:

- Ikke-sammenslåtte resultater (3 studier) ga utilstrekkelig bevis for testnøyaktighet i henhold til EGAPP-verktøyet, og kvaliteten på rapporteringen var dårlig.
- Enkeltstudier med relativt små prøvestørrelser, som alle brukte plasmaprøver, rapporterte meget god samsvar mellom de ulike testene (Cohens k: 0,80 til 0,86).
- For deteksjon av PIK3CA isolert sett er kostnadene for PCR mindre enn NGS. Imidlertid vil bruk av PCR for deteksjon av tillegg relevante mutasjoner øke totalkostnaden sammenlignet med NGS. For tiden er kapital og infrastruktur samt vedlikeholdskostnader høyere for NGS enn PCR.
- Gitt ca. 140 pasienter med metastatisk BC er kvalifisert for PIK3CA-mutasjonstesting i Norge årlig, ble kostnadene ca. NOK 322 000.
- Alle tester har fordeler og begrensninger, men p.g.a ufullstendig informasjon var en skikkelig sammenligning ikke mulig. Valget av en egnet test for deteksjon av PIK3CA-mutasjoner avhenger av tilgjengeligheten av test-metoder, økonomiske hensyn, prøvetype og risiko for falske negativer, og tidsbruk.
- Fremtidig forskning bør fokusere på å gjennomføre større kohortstudier med veldefinerte pasientpopulasjoner, som følger pasientene fra testing (eller ingen testing), gjennom behandling og sluttresultater. Robuste og replikerbare metoder, samt en standard sjekkliste for rapportering, bør brukes for økt klarhet.

Table of contents

Preface

This Health Technology Assessment (HTA) was commissioned by The National System for Managed Introduction of New Health Technologies within the Specialist Health Service in Norway (Nye Metoder). The following commission was given 26.08.2019, and updated 14.12.2020: "A single technology assessment of alpelisib for the treatment of HR+/HER2-, locally advanced or metastatic breast cancer is carried out by the Norwegian Medicines Agency (NoMA). The Norwegian Institute of Public Health is responsible for carrying out the assessment of the relevant diagnostic test." (ID2019_070). This HTA includes a summary and critical appraisal of studies reporting on accuracy, and feasibility of anaytical tests (Real time polymerase chain reaction, Next generation sequencing, Sanger sequencing, and Liquid chip technologies)for the detection of PIK3CA (Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha) mutations in patients with advanced or metastatic breast cancer, and a cost analysis of tests relevant for the Norwegian context (RT-PCR and NGS). The organisation of molecular tests services, the ethical, legal, and social impact (ELSI), and the patient preferences related to pharmaco-genomic testing have been addressed in a previous publication from NIPH (1). The aim of this report is to support well-informed decisions in health care that can lead to improved quality of services.

The internal project group included the following members affiliated with the Norwegian institute of Public Health:

- Gerd M Flodgren (GMF), Senior researcher, project leader, systematic review
- Vida Hamidi (VH), Senior researcher, health economy
- Jose Francisco Meneses Echavez (JFME), Researcher, systematic review
- Julia Bidonde (JB), Senior researcher, systematic review
- Gyri Hval (GH), Research librarian, literature search

We would like to thank our external experts for their contribution to the project: Tormod K Guren, Oslo University Hospital, Emilius AM Janssen, Stavanger University Hospital. We also wish to acknowledge research librarian Ingrid Harboe for peer reviewing the search strategy.

The authors and involved experts declared no conflicts of interest.

We emphasise that although the clinical experts have contributed with valuable input and comments, NIPH is solely responsible for the content of this report.

Kåre Birger Hagen
Director of Reviews and Health Technology Assessments

Kjetil Gundro Brurberg
Department director of Reviews and Health Technology Assessments

Gerd M Flodgren
Senior Researcher, and Project lead

Abbreviations

ABC	Advanced Breast Cancer
ALK	Anaplastic Lymphoma Kinase fusion oncogene
AIHTA	Austrian Institute for Health Technology
cfDNA	Cell-free DNA
CI	Confidence Interval
ctDNA	Circulating tumour DNA
DNA	Deoxyribonucleic acid
DRG	Diagnosis-related group
EGAPP	Evaluation of genomic applications in practice and prevention
EGFR	Epidermal Growth Factor Receptor proteins
ELSI	Ethical, Legal, and Social Implications
EMA	European Medicine Agency
ER	Estrogen Receptor
ESMO	European Society for Medical Oncology
EUnetHTA	European Network for Health Technology Assessment
FDA	American Food and Drug Administration
FFPD	Formalin-Fixed Paraffin Embedded samples
HER2	Human Epidermal growth hormone Receptor 2
HR	Hormone Receptor
HTA	Health Technology Assessment
INAHTA	International Network of Agencies for HTA
IQWIG	Institute for Quality and Efficiency in Health Care
KCE	Belgian Health Care Knowledge Centre
LOD	Limit Of Detection
MBC	Metastatic Breast Cancer
NGS	Next Generation Sequencing
NIHR	National Institute for Health Research
NIPH	Norwegian Institute of Public Health
NoMA	Norwegian Medicine's Agency

NSCLC	Non-Small Cell Lung Cancer
EMQN	European Molecular Genetics Quality Network
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
qPCR	quantitative Polymerase Chain Reaction (also called RT-PCR)
RET	REarrangement during Transfection (RET) oncogene
RHF	Regionale Helse Foretak (Norwegian)
ROS1 gene	Proto-oncogene tyrosine-protein kinase fusion protein
RTK	Receptor Tyrosine Kinase
RT-PCR	Real-Time Polymerase Chain Reaction (also called qPCR)
PD-LI	Programmed Death Ligand 1
PFS	Progression Free Survival
PICO	Population, Intervention, Comparison, Outcomes
SR	Systematic Review
STARD	Standards of Reporting of Diagnostic studies
TAT	Turn-Around time
TKI	Tyrosine Kinase Inhibitor
VAF	Variant Allele Frequency

Objectives

The main objective of this evaluation was to summarise available evidence on the analytical validity, the clinical validity, and the clinical utility of molecular tests (Sanger sequencing, PCR, NGS, and Liquid chip technology), for the detection of PIK3CA (Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha) mutations in men and postmenopausal women with hormone receptor positive (HR+)/human epidermal growth factor receptor2 negative (HER2–), locally advanced or metastatic breast cancer (BC) that has progressed during endocrine treatment.

More precisely we aimed to answer the following research questions:

- How accurately and reliably do each of these tests detect the biomarker in the laboratory (technical performance)?
- How accurately and reliably do each of these tests detect the biomarker in samples from patients with locally advanced or metastatic BC (e.g. tumour tissue, circulating cells, or cytology samples)?
- How well do each of these tests predict the effectiveness of treatment (e.g. shrinking of the tumour, or slowing down the disease process)?
- How well do each of these tests predict outcomes of importance to the patient (e.g. overall survival, and quality of life)?
- What are the potential adverse effects of using these tests to guide treatment decisions affecting patients?
- What are the advantages and limitations of these tests (i.e. the feasibility of tests in terms of biological tissue requirements, turnarond time, invasiveness, infrastructure, and training/expertise needed for running the analyses or interpreting the test results)

An additional aim was to estimate the costs related to testing for PIK3CA mutations in the same patient population and using analytical tests relevant for a Norwegian context.

Organisational aspects related to test services in Norway, ethical, legal, and social implications (ELSI), and patient preferences related to molecular pharmacogenomic testing have been addressed in a recent publication from NIPH (1).

Background

General background

Precision medicine is a term commonly used to describe treatments, including therapeutic agents, tailored to individual patients or groups of patients (2). The overall goal of precision medicine in oncology is to match pharmacological therapies to individuals to ensure that they receive effective treatment with minimal toxicity. This is particularly important for patients who may have a limited life expectancy (3).

A significant part of precision cancer medicine, involves the identification of a 'biomarker' associated with a particular cancer type, which typically constitute a unique mutated nucleic acid sequence, protein, glycoprotein, or group of proteins, expressed by the tumour cells but not normally by healthy cells (2). The presence or change in a particular biomarker is taken to predict which individual or group of individuals that are more likely to benefit (or not benefit) from a particular drug therapy (4).

The potential of precision medicine to improve patient outcomes, depend on the accurate identification of patients for treatment, through predictive biomarker testing (5). Hence, there is a natural dependency that exists between biomarker-based treatment and test. Unfortunately, designing a validated diagnostic assay to identify eligible patients for treatment does not guarantee accurate detection of the biomarker of interest and subsequent delivery of treatment (6). Effective use of biomarker tests and applying high-quality testing standards are fundamental to deliver precision medicine.

There has been a steady growth in the number of genomic tests available for use in healthcare services during the last two decades (7, 8). The task of determining the appropriateness of the plethora of different tests is a challenge for both clinicians and policy- and decision-makers (7). The clinical use of reliable tests to guide therapy selection depends on many related factors (i.e., analytical validation, clinical validation, specimen handling, reproducibility, information technology, and infrastructure), which all can affect the accuracy and reliability of test results and thereby the patient safety (9).

Condition/disease

Epidemiology

Breast cancer (BC) is the most frequently occurring cancer in women worldwide (10). In 2020, 3, 455 new BC cases were diagnosed in Norway, of which 3,424 were female and 31 were male (2).

A majority of BC (81%) are invasive, i.e. they spread from the walls of the glands or ducts where they originated and grow into surrounding breast tissue. BC is regarded as a group of diseases, consisting of at least 21 distinct histological subtypes, and four major molecular subtypes: (i) Luminal A hormone-receptor positive (HR+), and human epidermal growth factor negative (HER2-); (ii) Luminal B (HR+/HER2+), (iii) Basal like/Triple negative (HR-/HER2-), and (iv) HER2-enriched (HR-/HER2+) that differ in risk factors, presentation, response to treatment, and outcomes. The HR+/HER2- molecular subtype is most common and constitute around 73% of all BC cases (11).

The 5-year survival for patients with metastatic BC, i.e. patients with a cancer that has spread to other organs, at diagnosis is estimated to 34% in Norway (12). It has been suggested that 15 % of patients with BC will develop distant metastases within 3 years after diagnosis of the primary tumour (13). While BC diagnosed at an early stage (non-metastatic), may be cured in around 70–80% of cases, advanced or metastatic BC is currently considered incurable with available therapies (14). Treatment for patients with advanced disease therefore focus on interventions to prolong life and prevent or alleviate symptoms.

The PIK3CA (Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha) gene, is a subunit of the enzyme phosphatidylinositol 3-kinase (PI3K), which is important for many cell activities in the human body, including cell proliferation, migration of cells, production of new proteins, transport of materials within cells, and cell survival (15). Dysregulation of PI3K signaling is highly implicated in tumorigenesis (i.e. initial formation of a tumor), disease progression, and the development of resistance to standard of care treatments currently used in breast cancer patients (16). PIK3CA mutations may lead to increased activity in the PI3K enzyme and thus increased PI3K-Akt-mTORsignaling and endocrine resistance (15).

The most common PIK3CA alterations (69% of the total) have been reported to affect the kinase (exon 20, p.H1047R in 35% and p.H1047L in 4% of patients) and the helical (exon 9, p.E545K in 17% and p.E542K in 11% of patients) domains. Other PIK3CA mutation variants that are of unclear clinical importance are found in exon 4 (p.N345K), in exon 13 (p.E726K), and in exon 20 (p.G1049R)(17).

PIK3CA gene mutations are believed to be present in up to 36% of all primary BC cases, ranging from 9 to 45% depending of molecular subtype (18), and further to be most abundant in the HR+/HER2- molecular subtype (19). In this group may up to 36% of patients be expected to harbour PIK3CA mutations (20), which would translate into around 140 patients with metastatic BC in Norway who potentially might benefit from treatment with Pi3k inhibitors (e.g. alpelisib) annually.

Accurate and reliable detection of PIK3CA mutations is therefore important for appropriate treatment decisions i.e. to ensure that people who may benefit from treatment are correctly identified.

Progress, treatment, and care pathway

Most patients with HR+/HER2-, locally advanced or metastatic BC are offered endocrine therapy (ET). However, international data suggest that around 10-20% will develop resistance (21), while according to Norwegian data the proportion is around 30% (22). According to the ESMO consensus guideline from 2020 (23), a CDK4/6 inhibitor combined with ET is considered the standard of care for this patient group. The CDK4/6 inhibitor can be combined with an aromatase inhibitor (AI), or with an estrogen receptor antagonist (e.g. fulvestrant), in de novo or recurrent advanced BC, in first or second line and in cases of primary or secondary resistance (23). This is also in line with Norwegian recommendations (12).

PIK3CA mutations have been suggested to predict treatment effect of PI3K inhibitors (75) in patients with HR+/HER2- advanced or metastatic BC (12). Treatment with alpelisib plus fulvestrant has in a randomized, double-blind phase III study (SOLAR-1), been shown to result in significantly longer progression-free survival (PFS) as compared to placebo plus fulvestrant (median 11.0 months (95% CI: 7.5 - 14.5) versus 5.7 months (95% CI: 3.7 - 7.4) in 572 patients including postmenopausal women, and men, with HR+/HER2-, locally advanced or metastatic BC whose disease had progressed after receiving an aromatase inhibitor (24).

Thus it may become relevant to test cancer-cells for mutations in the PIK3CA gene, which if detected, may result in treatment with PI3K inhibitors (e.g. alpelisib) being offered (12). Alpelisib in combination with fulvestrant is approved for use in postmenopausal women, and men, with HR+/HER2-, PIK3CA-mutated, advanced or metastatic BC following progression on or after an ET-based regimen by the European Medicines Agency (EMA) (25), and the US Food and Drug Administration (FDA) (26), As of today, this drug is approved for use in Norway, but approval of reimbursement is under evaluation for this patient group.

Molecular tests for detection of PIK3CA mutations

There are four analytical methods for the detection of PIK3CA mutations described in the literature (20):

1) *Real time polymerase chain reaction (RT-PCR, or RT-qPCR),* is the most frequently used method for the identification of PIK3CA mutations (8). *RT-PCR* combine amplification and detection into a single step through the use of a variety of different fluorescent chemistries that correlate PCR product concentration to fluorescence intensity (27).
2) *Next generation sequencing (NGS),* or deep parallel sequencing, is a high-throughput method used to determine a portion of the nucleotide sequence of an individual's genome, which used DNA sequencing, by which multiple genes can be sequenced simultaneously (28).
3) *Sanger (direct) sequencing*, is also a DNA sequencing method, which involves electrophoresis and is based on the random incorporation of chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication.
4) *Liquid Chip Technologies*, or the mutant-enriched liquid chip method as described by Ming (29) consist of three steps: first a PCR amplification to eliminate the wild-type genes by restriction enzyme digestion, followed by selective amplification of the mutated DNA sequence, and hybridization of the mutated PCR product to a specific probe, which is precoated on the polystyrene microspheres and analyzed using an xMAP analyzer.

At present there are no available guidelines or recommendations regarding which diagnostic tests to use for the detection of PIK3CA gene mutations in advanced or metastatic BC.

Companion diagnostic tests

There are three companion tests for alpelisib (PIQRAY®) that are approved by the FDA for detection of PIK3CA mutations to select patients for drug treatment: FoundationOne®CDx (F1CDx), FoundationOne®Liquid CDx (Foundation medicine), and Therascreen PIK3CA RGQ Polymerase chain reaction (PCR) kit (QIAGEN. Manchester Ltd)(30). These tests enable the use of either tumour tissue specimens (e.g. DNA from FFPE) or circulating tumour DNA (ctDNA) extracted from plasma for the analysis. However, if the test is negative for PIK3CA mutations in plasma, the FDA recommends that patients should undergo testing for PIK3CA mutations in tumor tissue.

F1CDx is an NGS-based in vitro diagnostic device that uses DNA isolated from FFPE tumor tissue specimens for detection of substitutions, insertion, and deletion alterations, and copy number alterations in 324 genes and select gene rearrangements,

along with microsatellite instability and tumor mutational burden. F1CDx received FDA approval in December 2019.

In November 2020 FDA extended their approval to include FoundationOne®Liquid CDx, which used cell free DNA (cfDNA) from plasma and NGS with high-throughput hybridization-based capture technology to evaluate over 300 genes.

Therascreen PIK3CA RGQ PCR kit is used for qualitative detection of 11 mutations in the PIK3CA gene by real-time PCR, and uses DNA extracted from tumour tissue or from plasma (ctDNA) for the analysis. Therascreen received FDA approval in May 2019. This product, however, was recalled in 2021, due to high frequency of Q546R false mutation positive results (https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRes/res.cfm?id=185128).

Why is it important to conduct this assessment?

In this HTA we have summarised available evidence of the accuracy, advantages and limitations of tests relevant for the detection of PIK3CA mutations in men and post-menopausal women with HR+/HER2-, locally advanced or metastatic BC that have progressed under endocrine treatment. In addition, we have conducted an economic evaluation of the diagnostic methods relevant for the Norwegian clinical setting. This assessment was conducted to assist decision makers in making informed decisions regarding the delivery and organisation of molecular tests services in Norway.

Method

We have conducted a systematic review of the literature in response to this commission. The purpose of the molecular tests under study was prediction of treatment response or adverse events. We used a combination of the Evaluation of Genomic Applications in Practice (EGAPP) framework (31, 32), and the extended framework described by Pitini et al. to guide our assessment (33). A glossary is found in Appendix 1.

Literature search

Research librarian Gyri Hval (GH) developed the search strategy with input from the authors, and ran the electronic searches. Another research librarian (Ingrid Harboe) peer reviewed the search strategy. The search did not have any time limits or language restrictions. The full search strategy is provided in Appendix 2.

We systematically searched for literature from inception and up to October 2021, in the following electronic databases:

- Epistemonikos
- MEDLINE (Ovid)
- Embase (Ovid)
- Cochrane Database of Systematic Reviews
- Cochrane Central Register of Controlled Trials
- Epistemonikos
- ClinicalTrials.gov
- WHO ICTRP
- Current Controlled Trials

In addition we searched the HTAi Vortal, PROSPERO and the POP database in January 2020.

Inclusion criteria

We used the PICO (population, intervention, comparison and outcomes) framework to describe the inclusion criteria of the review (34). (Table 1) The exclusion criteria are described in Table 2.

Table 1. *PICO criteria for inclusion of studies*

PICO	
Population:	Men and postmenopausal women, with HR+/HER2- locally advanced or metastatic breast cancer (BC), that had progressed under endocrine treatment
Intervention (index test (s)):	• Sanger sequencing, • Polymerase Chain Reaction (PCR), • Next generation sequencing (NGS), and • Liquid Chip Technology, for the detection of PIK3CA mutations
Comparison (reference test(s)):	Head-to-head comparisons of the tests listed above
Outcomes:	• Analytic validity (sensitivity, specificity, assay robustness) • Clinical validity (sensitivity, specificity, predictive values, likelihood ratios), and concordance (i.e. agreement) • Clinical utility (e.g. overall survival, quality of life) • Advantages and limitations (e.g. amount and type of biological material needed, turnaround time, coverage, challenges related to the analysis or interpretation of test results)
Language:	English, Norwegian, Swedish, Danish, Icelandic, Spanish, and Persian
Study design:	Original studies (accuracy); Systematic reviews, and non-systematic reviews (feasibility)

Table 2. *Exclusion criteria*

Exclusion criteria

Study design:	Case reports, case series, conference abstracts, animal studies, and studies for which a full text article was not available
Population:	Studies solely including patients with other molecular BC subtypes, and studies with mixed populations in which results for advanced or metastatic BC were not reported separately
Intervention:	Other tests than those listed above

Outcome: Other outcomes than those listed above

Other exclusion criteria were studies that did not report a comparison between tests, or with cell-lines with known mutation status, or studies written in other languages than those listed above.

Selection of studies

We downloaded all titles and abstracts retrieved by the electronic searches into the reference management program EndNote (35) and removed duplicates. Two review authors (from GMF, JME, and JB) independently assessed the remaining titles and abstracts against the inclusion criteria using Rayyan (36). We obtained full text copies of potentially relevant studies, and assessed them in duplicate. We resolved disagreements by discussion. Reasons for exclusion of publications read in full text but subsequently excluded are reported in Appendix 3.

Data extraction and management

Two reviewers (GMF, JB and JFME) independently extracted data from each included study into a standardised and piloted data extraction form, which was adapted for use in this HTA. Any disagreements were resolved through discussion among review authors. We extracted the following data: citation, year of publication, setting, country, funding, conflicts of interest, study designs, language, and details on the PICOs:

- *Participants*: number, age, gender, race/ethnicity, socioeconomic status, time since diagnosis, previous treatment received, concomitant therapy/medication, etc.
- *Molecular tests*: technical details of tests, regulatory status, in-house or commercial test, previous tests conducted, sequence of tests if more than one test, test turnaround time, type and amount of biological tissue needed, etc.
- *Comparisons*: head-to-head-comparisons, intervention (index) test and comparator (reference tests) if applicable, or cell-lines with known mutation status
- *Outcomes*: analytical validity (sensitivity, specificity, assay robustness, quality control), clinical validity (i.e. sensitivity and specificity, positive and negative predictive values, and concordance), clinical utility (e.g. response rate to treatment, overall survival, quality of life), advantages and limitations of the different tests etc.

Quality of evidence and quality of reporting

Two authors (from GMF, JB, and JME) used the three-step process suggested by EGAPP working group (31) to assess the quality of the evidence of included original studies using what the EGAPP-group refers to as a 'chain of evidence'. See Appendix 4. The three steps include:

(i) Determining the hierarchy of the data source and study design (Level 1 to 4, of which level one is the highest) for the components of the evaluation (i.e. analytical validity, clinical validity, and clinical utility);
(ii) Determining the quality of individual studies (internal validity), and,
(iii) Grading the quality of evidence for the individual components of the chain of evidence (convincing, adequate, or insufficient).

We did not quality assess the included narrative review as these type of publications, by nature are considered to be low quality (37).

In addition, we used the Standards for Reporting of Diagnostic Accuracy (STARD) checklist (38), which contains 30 essential items for studies of diagnostic tests, to assess the quality of the reporting in the included original studies. One author (from GMF, JB, and JFME) assessed the quality of reporting, and a second author double-checked the accuracy of the assessment. Any discrepancies were solved though discussion between authors.

Compilation of results

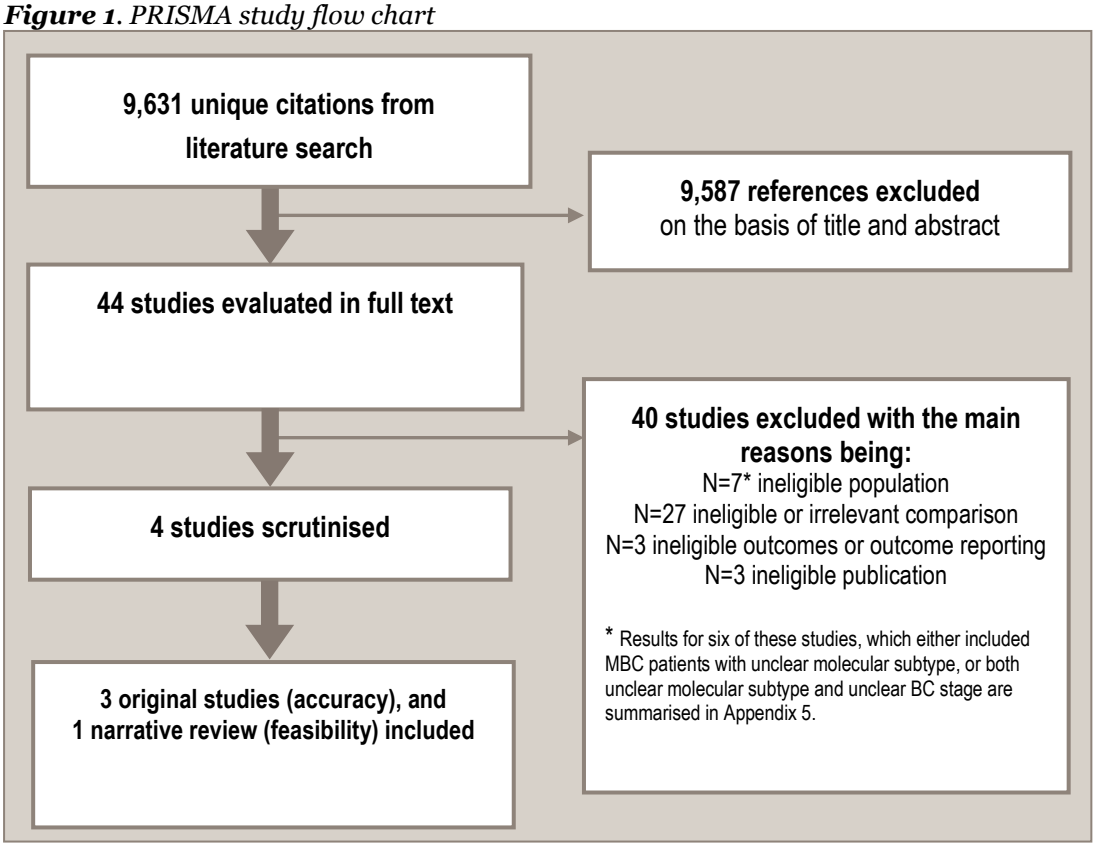
Meta-analysis was not feasible mainly due to heterogeneity in terms of intervention and comparator (reference) test used, and single studies providing evidence for each comparison. We have therefore provided a narrative summary of the available evidence from original studies on the concordance between different analytical techniques for the detection of PIK3CA mutations in text and tables. In addition, data on the advantages, and limitations of relevant tests retrieved from a recent review has also been summarised narratively. Additional data is provided in appendices.

Results

Search results

The databases search yielded 12,785 citations, of which 3,154 were duplicates. Of 9,631 unique citations 9,587 were irrelevant and directly excluded at title and abstract screening stage, leaving 44 citations to be retrieved in full text for further scrutiny. Three publications (39-41) that provided data for comparisons of tests (concordance) for the detection of PIK3CA mutations in ER+ and/or ER+/HER2-, advanced or metastatic BC, were considered relevant for this HTA. One narrative review that described advantages and limitations of relevant tests was also included (42). Studies read in full text but subsequently excluded are listed in Appendix 3, along with the reasons for exclusion. Figure 2 presents the PRISMA study flow chart (43).

***Figure 1**. PRISMA study flow chart*



Characteristics of included studies

See Table 3 Characteristics of included patients, Table 4 Overview of test comparisons and no of samples tested, Table 5 Characteristics of included studies.

Study design, and country of origin

We included three original studies (39-41). All studies had unclear study design (i.e. it was not clear whether data collection was planned before the intervention (index) and comparator (reference standard) tests were performed). One study was conducted in Australia (41), and one in the UK (39). In one study (44) the authors were from the UK, US, Germany, and Australia, but the actual setting of patient recruitment was not reported. The studies were published between 2019 and 2021.

Population: Selection of participants and their characteristics

Nteliopaulus and colleagues (39) included 50 unselected patients with radiologically confirmed metastatic BC (a total of 96 blood samples) attending the Breast Clinic at Charing Cross Hospital in London, UK. All patients were ER+ and resistant to endocrine therapy. The HER2 status was not reported. Time period for data collection/testing was not reported, and the included patients were sampled at different times throughout the course of their treatment. No further demographic information (e.g. age, gender, time since diagnosis) was provided. Thirty-five samples were tested with two different NGS platforms (*InVisionSeq™ ctDNA Assay and the Oncomine™ Breast cfDNA Assay*), but five of these samples were excluded due to variants unique to InVision.

O'Leary and colleagues (40) included 390 of 521 patients with ER+/ HER2- advanced BC enrolled in the phase 3 PALOMA-3 trial, who had paired baseline ctDNA analysis from two tests. Twenty-seven of the 390 samples were excluded due to various reasons (1 missing data, 20 samples not tested with all three ddPCR multiplexes, and six samples with no allele fraction estimated for ddPCR). All patients had progressed under endocrine therapy and were at the same treatment stage. No further characteristics of the included patients was provided. Three-hundred and sixty-three samples were tested with two different PCR techniques (*BEAMing and ddPRC*).

Zivanovich et al (41) included 234 consecutive women and men with newly diagnosed or previously established MBC of any histological subtype who were undergoing treatment. The cohort was representative for the clinical population and BC subtypes with around 78% being ER+/HER2-, (12% HER2+, and 10% triple negative BC). The median age of patients was 56 years. The majority of participants had newly diagnosed metastatic BC (n = 80; 34.2%) or disease progression following >2 prior lines of therapy (n = 89; 38%). The study population exhibited variable sites of metastatic disease, with the most frequent site being bone (n = 160; 68.4%). The median follow-up of participants was 15 months (range 1–46). One-hundred and

sixty-two samples were tested with both analytical techniques (*targeted NGS sequencing and ddPCR*).

***Table 3**. Characteristics of included patients with advanced(40) or metastatic (39, 41)BC*

Author Year	No of pts[a]	Age (median years)	Gender	BC subtypes	Treatment	Time since diagnosis	Other information
Nteliopaulus 2021 (39)	**50**	NR	NR	ER+	All were resistant to endocrine therapy.	NR	Unclear proportion of HER2- pts.
O'Leary 2019 (40)	**390 (27 excluded)**	NR	NR	ER+/HER2-	All had previously received endocrine therapy, and were at the same treatment stage.	NR	All enrolled in the phase 3 PALOMA-3 trial
Zivanovich 2020 (41)	**234**	All: 56 years (28 to 83) **ER+/HER2-**: 56 years (30 to 83)	All: Female: 232 (99.1%); Male; 2 (0.9%) **ER+/HER2-**:Female:180 (98.9%); Male:2 (1.1%)	ER+/HER2+:14 (6%); ER-/HER2+: 14 (6.0%); TNBC: 24(10.3%) **ER+/HER2-** : 182 (77,8%)	All: No treatment: 80 (34.2%) 1 or 2 lines:89 (38.0%) 3 or 4 lines:35 (15.0%) ≥5 lines: 30 (12.8%) **ER+/HER2-**: No treatment: 56 (30.8%) 1 or 2 lines: 70 (38.5%) 3 or 4 lines: 31 (17.0%) ≥5 lines: 25 (13.7%)	All:<1 year: 10 1-5 years: 107 6-13 years: 25 **ER+/HER2-** : <1 year: 72(39 1-5 years: 87 (6-13 years: 23	Any histological subtype undergoing treatment. Metastatic sites were mainly osseous (160/234) Unclear proportion of ER+/HER2- pts among the 162 included samples.

ER+: estrogen receptor positive; HER2-: human epidermal growth factor receptor 2 negative; MBC: metastatic breast cancer; NR: not reported; Pts: patients; a. Patients in the study by O'Leary had advanced BC, and patients in the two other studies had metastatic BC.

Intervention (index) test/Comparator (reference) tests

One study (39) compared different NGS platforms, one study (40) compared two PCR techniques, and one study compared ddPCR with NGS (41) for the detection of PIK3CA mutations. The authors typically did not refer to the tests as index or reference standard tests. We found no eligible study that compared Sanger sequencing, or Liquid Chip technologies with other tests for the detection of PIK3CA mutations in people with ER+/HER2-, advanced or metastatic BC. Nor did we find any comparative studies including any of the companion tests approved by the FDA for detection of PIK3CA mutations (i.e. FoundationOne®CDx (F1CDx), FoundationOne®Liquid CDx, and Therascreen PIK3CA RGQ Polymerase chain reaction (PCR) kit (30)).

The number of samples tested by two different analytical techniques ranged from 30 in Nteliopaulus et al. (39) to 363 in the study by O'Leary and colleagues (40) (median: 162 samples). In all three studies a sub-sample was tested with the second method. The reason for this was reported to be due to insufficient material for running the second test in one study (39)(Table 3).

***Table 4**. Test comparisons and no of samples analysed in included studies (N=3)*

Author Year	Total no of samples (no paired samples)	ddPCR (no samples)	BEAMing (no samples)	NGS- platform (a) (no samples)	NGS-platform (b) (no samples)
Nteliopaulus 2021 (39)	96 (35)			96 InVisionSeq™ ctDNA Assay	30[a] Oncomine™ Breast cfDNA Assay
O'Leary 2019 (40)	363 (363)	363	363		
Zivanovich 2021 (41)	234 (162)	234		162 Access Array™ system (Fluidigm)	

ctDNA: circulating tumour DNA; ddPCR: digital droplet Polymerase Chain Reaction; NGS: Next Generation Sequencing; Pts.: patients; a. Five samples were excluded due to being unique to the InVisionSeq™ctDNA assay.

Outcomes

All three studies reported concordance (Cohen's kappa) between tests (see Table 5). No study reported on the sensitivity or specificity of the tests. Only two of the studies (39, 40) provided a point estimate, with a measure of dispersion. Note that for one of the studies (41) we calculated the concordance between tests for PIK3CA detection, as results for detection of PIK3CA were not reported separately.
One of the included studies reported concordance for four PIK3CA mutation variants separately (40). The other two studies reported concordance for all detected PIK3CA gene mutations taken together.

Test and sample characteristics

Sample types and material requirements

All studies extracted DNA from plasma samples for the analysis. The amount of input material needed (i.e. plasma and/or- genomic DNA) ranged from 0.25 to 2 ml plasma, and from 1.3 to 20 ng total ctDNA plasma (See Table 4).

PCR techniques, and NGS platforms used

Two studies (40, 41) used the QX200 Droplet Digital PCR System (Bio-Rad Laboratories), and data analysis by the QuantaSoft package version 1.7.4.0917. O'Leary et al (40) initially screened samples with 3 multiplexes of custom Taqman ddPCR assays with FAM fluorophore mutant probes and HEX fluorophore wild-type probes, and performed a PIK3CA multiplex testing for p.E542K (c.1624GA), p.E545K (c.1633GA), p.H1047R (c.3140AG), and p.H1047L (c.3140AT). Zivanovich et al (41) used a multiplex assay targeting 20 hotspot somatic mutations in PIK3CA, ESR1, AKT1, and ERBB2 (of which four were in the PIK3CA gene: E545K, H1047L, H1047R, E542K). O'Leary et al (40) used BEAMing, which was performed by Sysmex Inostics, using the OncoBEAM assay, and quantified DNA with a LINE1 real-time quantitative PCR assay, after which DNA was subjected to PCR preamplification before the reaction products were split into variant-specific emulsion PCR and hybridization reactions.

Two studies (39, 41) used various NGS systems/platforms from different suppliers (Table 5). Nteliopaulus et al (39) did not describe what PIK3CA mutation variants that were covered by the two different NGS panels, but mentioned that five samples were excluded since they included variants that were unique to InVision seq.

Cut-off criteria/standard for test positivity

The cut-off criteria for test positivity for ddPCR varied somewhat across studies. Cut-off criteria /standards for positivity for NGS analyses was only reported for the Access Array™ system (Fluidigm) used in Zivanovich (41). One study (39) provided no cut-off criteria for either of the analytical methods used.

***Table 5**. Characteristics of included studies (N=3)*

	Nteliopaulus 2021 (39)	**O'Leary 2019 (40)**	**Zivanovich 2020 (41)**
Country	UK	Authors from UK, US, Germany and Australia	Australia
Aims	To assess the concordance and feasibility of InVisionSeq™ ctDNA Assay and the Oncomine™ Breast cfDNA Assay for the detection of mutations in plasma at low (<0.5%) variant allele fraction (VAF)	To assess the agreement in ctDNA mutation calling between BEAMing and ddPCR, two of the most commonly used digital PCR techniques for detecting mutations in ctDNA	To assess the feasibility and utility of applying routine comprehensive ctDNA profiling in mBC patients to guide clinical management.
Study design	unclear	unclear	unclear
Mutations/variants assessed	PIK3CA (and variants)[a,b]	:4 PIK3CA variants; p.E542K(c.1624GA), p.E545K (c.1633GA), p.H1047R(c.3140AG), and p.H1047L (c.3140AT)	ddPCR, multiplexed for 20 targetable hotspot mutations in four selected genes Access Array™ system: 39 genes recurrently mutated in mBC
Exons investigated	In Vision:2.8.10.21 Oncomine: 5,8,10, 21	-	-
Comparison	2 NGS platforms	2 PCR techniques	ddPCR vs. targeted NGS
Intervention	Oncomine™ Breast cfDNA	BEAMing	Access Array™ system (Fluidigm)
Comparator	InVisionSeq™ ctDNA Assay	ddPCR	ddPCR
Sample type	Plasma (cfDNA)	Plasma	Plasma
Material requirements	20ng total cfDNA	ddPCR: 0.25 mL equivalent or 1.3 ng, whichever was the greater BEAMIN: 2mL plasma	2ml plasma
Criteria for positivity	NR	*ddPCR* (multiplexed): positive results required at least 2 droplets on the mutant channel. *BEAMing*: positive calls were only made when allele fraction was consistent with >1 mutant copy being present in the initial reaction according to the results of the LINE1 quantitative PCR, with a threshold of 0.02% allele fraction from a negative control.	*ddPCR*: a sample was considered positive for a particular mutation (and consequently verified via singleplex assay) if the overlap between QuantaSoft images for the test sample and positive control showed detected droplets in the same cluster for the mutant DNA *Access Array™ system*: variants were only called if they were absent in both germline replicates but appeared in both technical replicates at >1% VAF (unless previously detected via ddPCR, in which case there was no restriction applied on VAF cutoff)
Outcomes reported	concordance	concordance (and changes in clinical care)	concordance

BEAMing: beads, emulsion, amplification, and magnetics; ddPCR: droplet digital PCR ; ER+: estrogen receptor positive ; HER2-: human epidermal growth factor receptor 2 negative; mBC: metastatic breast cancer; NGS: Next Generation Sequencing; NR: not reported; Pts: patients; PIK3CA: Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha; a. InVision assay profles a combination of SNVs and INDELs across 35 cancer genes (v1.4) or 36 genes (v1.5), as well as CNVs in 4 genes. b.The Oncomine assay analyses SNVs/INDELs covering>150 hotspots in 10 frequently mutated genes in BC. Seven genes were overlapping.

Results for concordance between tests

All three studies reported concordance between tests for the detection of PIK3CA mutations in samples from patients with ER+ and ER+/HER2-, advanced or metastatic BC. Concordance ranged between 0.80 and 0.87 across comparisons, with Cohen's kappa suggesting substantial to almost perfect agreement (Table 6).

In Nteliopaulus et al (39) three cases (10%) were discordant: two PIK3CA gene mutations were detected by Oncomine, but not by the InVisionSeq™ panel, and one mutation was detected by InVisionSeq™, but not by Oncomine.

In O'Leary and colleagues (40) 18 cases (~5%) were discordant between tests: three PIK3CA mutations were detected by ddPCR but not by BEAMing, and 15 mutations were detected by BEAMing but not by ddPCR. The composite allele fractions in the discordant cases were according to the authors often close to the calling threshold of 0.1% allele fraction.

In Zivanovich et al (41) eight cases (~5%) were discordant between tests: six PIK3CA mutations that were undetected by targeted NGS were detected by ddPCR, and in two samples were ddPCR inconclusive for PIK3CA mutations, but detected (positive) with NGS. The authors suggested that mutations detectable with ddPCR might have been below the analytical sensitivity of the assay, (i.e.<1%VAF) and therefore undetectable with targeted NGS.

Results for concordance between tests for PIK3CA gene mutation variants, which was reported in one study (40) comparing two PCR techniques, suggest lower concordance for less prevalent PIK3CA mutation variants, with κ ranging from 0.44 to 0.90 (data not in table).

***Table 6**. Results for comparisons between tests: concordance*

Author Year	Mutation/variant	Comparison	Total no of samples tested with both methods	No of concordant cases	No of discordant cases	Concordance–Cohen's kappa
Nteliopaulus 2021(39)	PIK3CA	NGS (Oncomine™ Breast cfDNA Assay) vs. NGS (InVisionSeq™ ctDNA Assay)	30	27 (13 pos, and 14 neg.)	3	0.80 (95%CI, 0.59 to 1.00)
O'Leary 2019 (40)	PIK3CA	PCR (BEAMing) vs. PCR (ddPCR)	363	345 (80 pos, and 265 neg.)	18	0.87 (95% CI, 0.81 to 0.93)
Zivanovich 2021 (41)	PIK3CA	PCR (ddPCR)vs. NGS (Access Array™ system (Fluidigm))	162	162 (30 pos., and 124 neg.)	8	0.85

BEAMing: Bead-Based digital PCR in emulsion; cfDNA: cell-free DNA ; ctDNA: circulating tumour DNA; ddPCR: digital droplet Polymerase Chain Reaction; PIK3CA: Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha

Quality of included evidence: results of the EGAPP tool

See Appendix 6. Quality of included evidence: results of the EGAPP tool assessment. The EGAPP tool is described in Appendix 4.

The results of the 3-steps EGAPP quality assessment are described below. The tool, as mentioned earlier, involves assessment of analytical validity, clinical validity and clinical utility. None of the included studies had as the main aim to compare the tests analytical validity in the laboratory (technical performance), nor did any study report data on clinical utility (risks and benefits for the patients). We have summarised results related to clinical validity below.

***Step 1**. Assessment of the study hierarchy (level 1-4, with 1 being the highest):* Step
We judged one (39) of the three included studies to be level 3 in the hierarchy of study evidence proposed by EGAPP (lower quality case-control or cross-sectional studies). Two studies (40, 41) were judged to be level 1 studies (well-designed cohort studies).

***Step 2**. Assessment of the internal validity of included studies:*
Clear description of the disorder/phenotype and outcomes of interest. Status (biomarker positive or negative). In the three studies a subgroup of cases (e.g. samples with sufficient material, or samples with variants that were included by both tests) were assessed with both the intervention and the comparator test.

Adequate description of study design and test/methodology. Two of the three studies provided information on study design (40, 41). All studies described the tests under study, but it was unclear if the level of detail was sufficient to allow replication. Two of the studies (40, 41) however, provided quite extensive descriptions of the tests and methods used.

Adequate description of the study population. Two studies (40, 41) provided adequate inclusion criteria (but no exclusion criteria). The included participants were described in more detail in two studies (39, 41), while the third study (39) provided little or no information.

I*ndependent blind comparison with appropriate, credible reference standard(s).* None of the studies referred to the tests as index or reference (standard) tests. One study (41) reported that they removed samples from the analysis in order to restrict the analysis to high confidence variants and control for sequencing and/or PCR artefacts, if at least 1 replicate had a read coverage depth <100 in either germline or plasma samples (three samples were removed). None of the studies reported how indeterminate results were handled. One study (40) reported that the analyses were blinded and conducted at independent laboratories.

Analysis of data. Possible biases were identified, and their potential impact were at least partly discussed in all studies. Point estimates of sensitivity and specificity were not provided in any of the studies. All studies reported on concordance, and in two

studies was a measure of dispersion provided (39, 40). No study reported estimates of positive and negative predictive value.

Step 3. *Grading of the certainty of evidence (convincing- adequate- inadequate)*
We judged that all three included studies provided inadequate evidence for test accuracy/clinical validity.

Quality of reporting–results of the STARD checklist

The results of the STARD checklist revealed relatively poor reporting across included studies, but with more information provided in two studies (40, 41). The results of the 30 items in the checklist are summarised in the text below and presented in detail in Appendix 7.

Item 1-4. Title, abstract, background. None of the three included studies self-identified as a test accuracy study in the title, but that accuracy related outcomes were of interest was indicated in the abstract. All three provided a structured abstract (39-41), a scientific and clinical background, with a description of relevant tests, and study objectives (but no hypothesis).
Item 5. Study design. The study design in the three studies was unclear
Item 6-9. Methods-participants. Two of the three studies (40, 41) provided adequate inclusion criteria. In these studies patients were included based on participation in clinical trials. In one study the HER2 receptor status was unclear (39). Two studies (39, 41) provided information on location and timing of patient recruitment. One study (39) described patients as 'unselected', and in another study as 'consecutive' (41). One study (40) derived patient included in the PALOMA-3 trial but it was unclear if participants formed a consecutive, random or convenience series.
Item 10-13. Test methods. While the tests were described to some degree in all three studies, not all three phases (pre-analytical, analytical, post-analytical) were clearly described, and it was unclear whether the detail provided was sufficient to allow replication. Two studies (40, 41) provided a more comprehensive description of the tests. No study provided a rationale for choosing the comparator test, and one study (41) stated that a reference standard for ctDNA analyses was lacking. Two studies (40, 41) described the cut-off criteria for positivity for both tests, while the third study did not provide a cut-off criteria for either. One study (40) reported that the analyses were blinded and conducted at separate laboratories, while the other two studies provided no information on blinding.
Item 14-18. Analysis. All studies provided information on the methods for estimating or comparing measures of diagnostic accuracy. No study provided information on how indeterminate results were handled. One study (40) included exploratory analyses of variability related to test accuracy. No study provided any information on the intended sample size.
Item 19-22. Participants: Two studies used a patient flow diagram (39, 41). One of the three studies reported baseline demographics and clinical characteristics of

participants (41). Zivanovich (41) also reported that the tests were run in parallel, and thus no clinical intervention could have been administered in between tests. O'Leary (40) reported that the tests were 'concurrent paired'. The third study(39) provided no information on the timing of tests.

Item 23-25. Test results: All three studies cross-tabulated cases against the results of the comparator test and provided results for concordance between tests. Two studies provided a measure of dispersion (39, 40). No study reported on adverse events related to the testing.

Item 26-27. Discussion. The three included studies discussed study limitations, biases and implications for practice to a various degree.

Item 28-30. Other information. All studies referred to a study registration/protocol (or a trial), but none to a study registration/published study protocol. The role of the funders was reported in one (41) of the three studies. See Appendix 8 for details on financial support, role of funders, and conflicts of interest.

Advantages and limitations of tests for detection of PIK3CA

Relevant tests

One narrative review (42) provided information on advantages and limitations of RT-PCR, NGS, and Sanger sequencing for the detections of PIK3CA gene mutations in ER+/HER2-, advanced or metastatic BC. The review provided no information on the liquid chip methodology. The results of this review, and information from personal communication with experts, are summarised below.

Type of samples, material requirements, and coverage

Both DNA from tumour tissue specimens (e.g. FFPE) and DNA extracted from plasma (ctDNA) may be used for the detection of PIK3CA gene mutations. There are advantages and limitations with both methods. The main advantages with liquid biopsies, as compared to tissue biopsy, as suggested by Fusco et al are their noninvasiveness and repeatability, allowing for a timely follow-up of treatment response and disease progression, and potentially overcoming spatial and temporal heterogeneity of the tumor (42). A major limitation of liquid biopsies however, is the lower concentration of tumor-derived DNA compared to tumour tissue specimens. According to the review by Fusco (42), liquid biopsies should not be used for analysis with direct (Sanger) sequencing, and there is a chance of fails and false-negative results when tissue samples are used. RT-PCR is very sensitive when liquid samples are used, but when tissue samples are used the risk of fails and false-negative results increases. NGS however, shows high sensitivity when tumour tissue specimens are used for the analyses, but a greater risk of fails and false-negative results when liquid samples are used. It should be noted that reflex testing using tumour tissue specimen is recommended if liquid samples test negative for PIK3CA using NGS or RT-PCR (45).

No information on the amount and quality of input material required for the different analytical techniques was provided in the review. Nor was any information regarding the cut-off criteria/standards for positivity of tests (42).

RT-PCR assays may be multiplexed but will still only cover a limited number of mutations. The technique uses 96 or 384-well plates by which many different RT-PCR products or many different samples can be analysed in one run. The technique is cheap and widely available, but the method detects mostly known mutations (46). NGS however, is easily multiplexed for numerous mutations (thereby avoiding the need for sequential testing) and has the capacity to detect both known and unknown mutations, which is two major advantages with this method. Sanger sequencing, on the other hand is not that easy to multiplex, and it is not possible to combine many different patients for the same sequencing reaction, which easily can be done with NGS (46). A Sanger Instrument might have between 4-16 capillaries, and a capacity of 4-16 reactions using different dyes. Some Sanger systems allows combinations of different dyes, which enables testing of up to 32 samples for one gene. An advantage of Sanger sequencing is the read length which is much longer than for NGS (46).
No information regarding the coverage of PIK3CA gene mutation variants across different PCR and NGS assays was provided in the review (42).

Turnaround time (TAT)
The total TAT is the interval between when a test is requested to the time a treatment decision is made (56). While evidently some of the steps are the same for the different tests, the time requirement for the analysis, interpretation of results, and communicating the results to the treating physician in an accessible form, may differ. TAT is in addition dependent on whether the test is 'in-house', or if analyses are centralised to certain hospitals, or to facilities outside the hospital. PCR has a short TAT which typically is between 4 hours and 1 day, while NGS may require around 3 days, but in some cases up to 1-3 weeks (46). No information was available on TAT for Sanger sequencing.

Facilities and expertise needed for analysis/ interpretation of results
NGS is a complex process, which requires high level of infrastructure and bioinformatics capability. Special expertise is also required for analysis with PCR. Time and expertise required for interpreting the results of different tests, and communicating these to the treating physician, was not mentioned in the Fusco review (42), nor was the use of 'data analysis pipelines' for NGS analyses.

PIK3CA testing at Norwegian hospital trusts
Not many Norwegian hospitals have Sanger sequencing any longer, instead most hospitals, especially the non-university hospitals, went directly from quantitative PCR to using NGS (46). Liquid chip methodologies are not used for diagnostic purposes in Norway (46).

Tissue samples are exclusively used for diagnostic purposes and treatment decisions at Norwegian hospital trusts, while plasma samples may be considered for patient follow-up after treatment (46).

Laboratories at Norwegian hospitals typically run NGS only one time/week (46), although the use of NGS for cancer diagnostics is now rapidly increasing (29). For the NGS platforms currently in use in Norway the TAT varies from 7-12 days (46). The TAT depends on the fact that DNA/RNA are isolated only certain days a week, and that the laboratories only run NGS one day a week. This is done to optimize and most efficiently use the reagents, kits, chips, and flow-cells, by running samples from different patients in parallel. Newer NGS systems from both Illumina and Thermo Fisher will be able to deliver a one-day workflow (24 h) i.e. here referring to the time from retrieval of DNA to when the analysis is done for some gene panel assays. For some NGS systems the reagents used for one sample will be the same as for 10 samples, and the flow-cell may be reused (46).

Results from a Norwegian survey suggest variations in the reporting of results of NGS analyses across trusts, a lack of reporting guidelines, as well as guidelines for deciding which gene panels to use, and which genes to test for (47). All Norwegian hospital trusts that use NGS, also use bioinformatic tools that are integrated in the platforms, various sky-based solutions, and free of charge data bases to help with interpretation of NGS results. However, while this use varies locally, it is not clear how these tools compare (47).

Economic Evaluation

General

The health care sector, similarly, to society in general, is restricted by limited resources and budget constraints. In Norway, health service interventions are to be evaluated against three prioritization criteria: the benefit criterion (increased longevity and/or improved health-related quality of life), the resource criterion, and the severity criterion (absolute shortfall) (48). Norwegian policy documents indicate that the priority-setting criteria are to be evaluated together and weighted against each other. This is to be done by means of a health economic evaluation.

Health economic evaluations are important tools for decision makers facing questions on how to prioritise health technologies and maximise health benefits using limited resources. The basic aim of any economic evaluation is to identify, measure and compare health consequences and costs of the alternatives under consideration in an incremental analysis, one in which the differences in costs are compared with differences in health consequences.

Identifying the place of a molecular test within care pathways is crucial, not only to guide the selection of a relevant comparator, but also to guide the use of the companion drug and subsequent treatment pathways to be modelled. The exact place along the treatment pathway where testing occurs may change the cost-effectiveness of the intervention because of differences in the type of treatment subsequently received and the costs and outcomes arising from these (49). In the published health economic evaluations of precision medicine, the cost of molecular testing and the accuracy of a test were reported as important factors, which influence the cost-effectiveness of targeted interventions (50, 51).

Ideally, diagnostic interventions should be supported by studies that follow patients from testing via treatment to final clinical outcome, so-called end-to-end studies (52). This combination of data will enable decision-makers to evaluate the overall cost-effectiveness of using a particular test-drug combination versus not using the drug at all or using the drug without the test (53). In other words, this will enable decision-makers to evaluate the predictive value of the test and the relevant biomarker and assess the health outcomes and economic consequences of using genomic test and the consecutive treatment in combination.

Key data needed for an economic evaluation of a molecular test include outcome data on the clinical effectiveness and utility of the technology, changes in health status as well as resource use and related costs of the affected patient population and the uptake of the test. Fundamentally, the challenges relating to the data requirements for the economic analysis of a pharmacogenetic intervention revolve around the availability and quality of existing data (49).

Method

In this health technology assessment, we have in line with the commission performed a simple cost analysis. We have in collaboration with the experts from the Norwegian regional health authorities, estimated the costs associated with diagnostic test relevant for Norwegian clinical settings (i.e., PCR and NGS), based on the micro-costing method. Micro-costing is a highly detailed health economic costing approach in which all the underlying resources required for an intervention or activity, such as equipment, consumables, and staff time are identified, and then unit costs are attached to this resource used to generate an overall cost (54).

PIK3CA testing can be performed on either tumour tissue or liquid (plasma) samples. Where a tissue sample (e.g., based on the site of local recurrence or metastasis) is not readily available, the test could potentially be performed on a plasma sample. However, test accuracy is typically lower when liquid samples are used, as compared to tumour tissue (42), due to the lower concentration of tumor-derived DNA in plasma. No identification of a PIK3CA mutation using the plasma test would result in patients needing a tissue biopsy for further consideration for treatment with alpelisib in combination with fulvestrant, due to the higher rate of false negatives using the plasma compared with the tissue test (45). It seems therefore reasonable to assume that the costs associated with PIK3CA testing using liquid samples may be higher than testing using the tissue samples. However, due to uncertainty associated with data on resource use, and as we did not identify any studies that reported the sensitivity and specificity of tests for the detection of PIK3CA mutations in liquid samples, our micro-cost analysis is limited to the cost associated with RT-PCR or NGS testing using tissue samples. It should be noted that according to our experts, tissue samples are used for all diagnostics at Norwegian hospital trusts, and liquid samples are used only for patient follow-up (46).

Implementing precision medicine in oncology is potentially a costly investment and it requires testing multiple patients to identify a specific group of responders to a targeted treatment. Currently, more and more diagnostic tests and precision medicines for particular diseases become available (55). In most solid tumour cancers, a set of parallel tests are to be performed on a number of molecular biomarkers to decide between a range of precision medicines (55). Therefore, we also considered the multigene testing and testing samples from multiple patients in the estimation of the costs related to the diagnostic methods.

Furthermore, we calculated the costs per test associated with the relevant diagnostic methods both using tissue and plasma samples based on the current Norwegian tariff rates for clinical laboratory services. A comparison between the estimated costs based on the micro-costing method and the costs calculated based on the tariff rates may assist the decision makers to make decision regarding the revision of the current tariff rates relevant for using molecular tests. We expressed relevant costs in 2022 Norwegian kroner (NOK).

Eligible population for detection of PIK3CA mutations

In 2020, 3,455 new BC cases were diagnosed in Norway, of which 3,424 were female and 31 were male (56). The average age at diagnosis in the last five years is 62 years for BC in Norway (57). Approximately 80% of Norwegian patients with BC belong to the HR+/HER2- molecular subtype (Communication with professor Bjørn Naume, OUH). The actual number of advanced BC patients in Norway is currently not reported in national registries. According to expert opinion 420 to 500 Norwegian patients annually have metastases at diagnosis or will develop metastases. It is reported that around 30% will progress under ET in Norway (22). Based on these assumptions around 138 patients would be eligible for PIK3CA testing per year in Norway (Table 7). Of the patients eligible for testing around 35% are believed to harbour PIC3CA mutations (20), i.e., around 48 patients.

Table 7. *Estimated number of patients eligible for detection of PIK3CA mutations*

	Number of patients	Comments
New BC cases	3,455	Cancer Registry of Norway 2020 (56)
HR+/HER2-	2,764	About 80% (expert opinion)
Locally advanced or metastatic	460	About 420-500 (expert opinion)
Progressed during ET	138	About 30% (22)
Candidate for PIK3CA testing	**≈140**	

BC: breast cancer; ER+: estrogen receptor positive; ET: endocrine treatment; HER2-: human epidermal growth factor receptor 2 negative

Results

This section presents the costs associated with different diagnostic methods based on a micro-costing analysis and based on current Norwegian reimbursement rates for clinical laboratory services.

Costs associated with different diagnostic methods based on data of micro-costing analysis received from the Norwegian hospitals

The costs associated with the relevant diagnostic methods for the detection of PIK3CA mutations in a Norwegian context are calculated based on the estimated resource use related to the use of the diagnostic tests for the detection of the relevant biomarkers in advanced lung cancer presented in our previous publications (1, 58). According to our experts, these data are also relevant for estimating the cost for PIK3CA testing.

To identify and measure the resource use and costs associated with the relevant diagnostic methods, we contacted four Norwegian regional health authorities. We received information from three University hospitals: Stavanger University Hospital, St. Olav's University Hospital, and Oslo University Hospital. The estimated costs vary according to the approach used to estimate costs by different hospitals, however, all these estimates included direct costs such as consumables and supplies costs and the cost associated with staff time. The resource use and unit costs are presented in Table 8. It should be noted that the costs for NGS are estimated for a panel that can analyse six (and two control) samples and six biomarkers simultaneously. The costs for RT-PCR were estimated for testing one biomarker.

Due to data consistency, our analyses are based on information received from Stavanger University Hospital (46). Data sent from St. Olav's University Hospital and Oslo University Hospital are presented in Appendix 9.

The costs associated with using RT-PCR for testing one biomarker were estimated to be around NOK 2,800 per patient. If the RT-PCR method will be used for the detection of multiple gene mutations in patients with advanced BC (e.g., EGFR/PIK3CA and BRAF), different tests would be run, and consequently, the reagent costs for the detection of one biomarker (varied between NOK 1,400 to NOK 2,500 per patient) should be multiplied by the number of gene mutations to be detected. Regarding personal costs, it is expected that the costs would be decreased as less time per patient would be spent for testing several patients at the same time. However, as the main costs are related to reagent costs, the total costs for testing several gene mutations using PCR are expected to be higher than NGS. In addition, it should be noticed that RT-PCR is a specific technique, and it lacks somewhat in sensitivity. It has been reported that the use of RT-PCR analysis using tissue samples would increase the risk of false negative results (42).

Table 8. *Estimated cost associated with using different diagnostic methods*

Diagnostic methods	**RT-PCR ***	**NGS ****	
	1 patient NOK	1 patient * NOK	10 patients NOK
Reagent costs***	1,380	14,480	1,730
Personal costs	1,459 (Bioengineer, Molecular biologist, pathologist)	1,821 (Bioengineer, Molecular biologist, pathologist)	594
Sum	**2,840**	**16,300**	**2,330**

Source: (46). The numbers are rounded.

** The estimated costs are associated with testing one sample.*

*** The costs are estimated for Oncomine Focus panel. The panel can analyse 8 samples and 6 biomarkers simultaneously.*

**** The reagent costs are varied between NOK 1,400-NOK 2,500 per patient, depend on the type of the commercial kit.*

The results of the cost analyses showed that NGS can be an affordable diagnostic method. The method is considerably more expensive if only one patient (one sample) would be tested. However, the foremost advantage of NGS technologies is the massively parallel sequencing capability. It means sequencing of multiple targeted genomic regions from the multiple samples in the same run. As the results showed in Table 8, the cost associated with NGS testing will be significantly decreased (approximately NOK 2,300 per patient) when parallel tests are performed on several biomarkers from multiple patients.

Based on this background, and feedback from experts that discovery of other mutations in the PIK3CA signaling pathway, and /or mutations that predict resistance, will make broader testing with NGS a reasonable choice for patients with metastatic BC. We have therefore estimated the costs associated with using NGS for the detection of PIK3CA along with other biomarkers to be approximately NOK 322,000 in about 140 eligible patients per year in Norway.

Preparing the biopsy

According to our experts, the costs associated with preparing the biopsy are the same for all diagnostic methods. Therefore, the biopsy costs are not included in the comparison between diagnostic methods. Based on data we have received from St. Olav's University Hospital the cost per biopsy per patient is approximately NOK 270 (Table 9).

Table 9. *Biopsy costs per patient*

	Costs per biopsy (NOK)	Comments
Formaldehyde 20 ml	12.66	50 pieces per package
Biopsy forceps	163.50	20 pieces per package
Biopsy wraps	1.24	Bio-wraps (100 pieces per package)
Personnel costs	92.00	10 minutes per biopsy
Total cost per biopsy per patient	**269.40**	

Source: St. Olav's University hospital (personal communication)

The costs related to the infrastructure, quality assurance and maintenance

PCR

Most university hospitals and some of the regional hospitals can perform PCR in Norway. The infrastructure costs for PCR are estimated to be around NOK <650,000, including the cost associated with qualification of gene expression (using Tapestation system) (46).

The maintenance costs of PCR instrument (incl. calibration, annual control and quality samples control) are calculated to be 25,000 per year. External quality rounds (NordiQc, EMQN) are estimated to cost around NOK 5,000 per biomarker/per year (46).

NGS

Currently, about eleven pathology departments in Norway have NGS equipment. Depending on the equipment capacity, it is estimated that the costs related to equipment and supplies investments are approximately 3 to 4 million NOK. The maintenance costs of NGS instruments are calculated to be about NOK 150,000 per year.

Based on the information from the experts, the validation process takes a lot of time and different types of expertise (including bioengineers, pathologists, engineers) are involved in this process. At least one NGS kit is used for the validation process. The price of such a kit varies from NOK 30,000 to 80,000. The validation process is only

performed once when the method is established. After that, external quality control system is used to check that all is still functioning adequately.

Costs associated with different diagnostic methods based on current reimbursement scheme for clinical laboratory services

In Norway, outpatient laboratory services are financed on block grants, co-payments from the patients, and tariff rates for clinical laboratory services (59) while laboratory services for inpatients are financed on case-based payment (Diagnosis-related group, DRG).

A new reimbursement scheme for outpatient pathology is recently introduced by the health Economics Administration (Helfo) (60). In the new reimbursement scheme, activity codes (Norwegian Pathology Code (APAT)) are used instead of the tariff rates. The new reimbursement scheme was implemented first at the Western Norway Regional Health Authority on the 1st of January 2022. It is planned that the new scheme will be implemented in all regional health authorities by 2023 (60).

In Table 10, we present the costs per investigation associated with different diagnostic methods used for the detection of PIK3CA mutations based on the new reimbursement scheme. The average cost is calculated by multiplying the reimbursement rate by two (61).

Table 10. *Average costs associated with using the relevant diagnostic methods based on the reimbursement scheme**

Diagnostic methods	Costs per test (NOK)	Comments
RT-PCR	368	PATA2 and PATA3
NGS	14,544	PATA2 and PATA3
Biopsy	1,372-1,786	PATP4 or PTAP5

Source: Lovdata poliklinikk-takster 2022 (59)

The reimbursement rate is only used for outpatients or samples sent to the laboratory.

Based on the new reimbursement scheme, the costs relevant to RT-PCR using tissue biopsies are estimated to be approximately NOK 1,950. This reimbursement rate will be NOK 368 if RT-PCR will be used for the detection of PIK3CA mutations in liquid samples. It is important to note that in case of negative results by using liquid samples, i.e., no identification of a PIK3CA mutation, a tissue biopsy should be also analysed to avoid false-negative results. This would result in higher costs associated with PCR testing using liquid samples.

A comparison between the estimated costs based on the micro-costing method and the costs estimated based on the current reimbursement scheme indicated that the current reimbursement rates are generally insufficient to cover the costs of running RT-PCR. Regarding NGS testing, depending on the number of genes analysed, different reimbursement codes can be used. In our analysis, we used the relevant code for testing 1-100 genes which can cover the costs of running the small NGS assay (for example, using Oncomine Focus assay consists of 35 hotspot genes). For the expanded gene panels which require a thorough pathology assessment and a significantly more advanced data analysis, and used mainly for the experimental treatment, the reimbursement rates for testing 101-400 genes and over 400 genes (without the costs related to biopsy) are estimated to be approximately NOK 24,900 and NOK 35,700, respectively (59).

Discussion

Summary of main results

We included three original studies that reported concordance between tests for detection of PIK3CA mutations in liquid samples from patients with locally advanced or metastatic BC, of which a majority belonged to the ER+/HER2- BC subtype (39-41). The median number of samples tested in the included studies was 162 (range: 30 to 363). No study reported test sensitivity or specificity, or adverse events related to the testing.

Un-pooled results from these three studies, of which one study compared two different PCR tests (40), one two different NGS platforms (39), and one study compared ddPCR with NGS (41), suggest good concordance (κ=0.8 to 0.86) between the tests for detection of PIK3CA mutations in plasma. The results further suggest (40) lower concordance for more rare, less prevalent PIK3CA mutation variants (κ range: 40% to 90%). However, the quality of the included evidence from the three studies was overall inadequate, and reporting was poor. We did not identify any studies evaluating the accuracy of tests for detection of PIK3CA mutations in tumour tissue samples. Neither did we find any eligible studies of Sanger sequencing or Liquid chip technologies, or any systematic reviews/meta-analyses on the accuracy of tests for the detection of PIK3CA mutations in our population of interest. It should also be noted that none of the companion tests approved by the FDA for the drug alpelisib were evaluated in the included studies. This is maybe not so surprising since the average post-launch delay in clinical uptake of a companion test has been estimated to 4.5 years (6).

Like in our previous reports (1, 58) we did not identify any end-to-end test accuracy studies with data on clinical utility, i.e., none of the included studies reported on outcomes of importance to patients (e.g., overall survival, quality of life), or how well the test(s) could predict the treatment effects (e.g., shrinking of the tumour or slowing down of the disease progress).

In summary, the evidence-base for the diagnostic accuracy of molecular tests for detection of PIK3CA mutations in men and postmenopausal women with HR+/HER2-, locally advanced or metastatic BC is scarce and incomplete.

Quality of included evidence and quality of reporting

The quality of the included evidence for the accuracy of tests from the three studies was overall inadequate. None of the included studies reported accuracy statistics (such as sensitivity and specificity), which typically is reported in diagnostic accuracy studies (62), but reported concordance, even if there are other, possible more robust, statistical methods that can be used when no credible reference standard is available (63). The quality of reporting was poor.

Feasibility of tests

Comprehensive information on the advantages and limitations of tests relevant for the detection of PIK3CA mutations in patients with ER+/HER2- locally advanced or metastatic BC, was not easily found in the literature. We identified only one relevant narrative review, which provided limited information on RT-PCR, NGS and Sanger sequencing, and no information on Liquid chip technologies. No fact sheets/ recommendations from ESMO regarding which analytical technique to use for the detection of PIK3CA mutations in liquid or tumour tissue samples from patients with advanced or metastatic BC were available. All analytical techniques appear to have both advantages and limitations, but since the information from the one included narrative review was far from complete an adequate comparison of tests was not feasible. However, a big advantage of NGS, over the other methods, is that it is easily multiplexed, and allows the detection of both known and unknown mutations. We however, found no information regarding the actual sensitivity and specificity of NGS in detecting PIK3CA mutations in liquid or tumour tissue biopsies from patients with ER+/HER2- locally advanced or metastatic BC.

Recommendations from ESMO published in 2020 (64) regarding the use of NGS-panels in metastatic diseases, do not recommend the routine use of NGS for patients with metastatic BC. They state, and I quote: "*Considering that somatic sequencing cannot fully substitute germline BRCA testing, that PIK3CA status can be determined by PCR on the three hotspots and pending that HER2 testing is accurately done by immunohistochemistry (IHC) in the local centre, there is currently no need to perform tumour multigene NGS for patients with MBC in the context of daily practice.*" However, discovery of other mutations in the PIK3CA signaling pathway that also trigger treatment with PIK3CA inhibitors, and /or mutations that predict resistance, will most likely make broader testing with NGS a reasonable choice for patients with metastatic BC (communication with experts).

Strengths and limitations

The HTA:

We developed a robust search strategy and conducted a comprehensive search for studies of relevance for our research questions. Screening, data extraction and quality assessment, were all done in duplicate to minimise bias and reducing the risk of missing important evidence. This is the third pilot of our assessment framework for pharmacogenomic tests, which has been piloted in two previous publications from NIPH ((1, 65).

Results for test concordance are reported narratively as meta-analysis was not feasible due to few included studies, and only single studies providing data for each comparison.

The included studies were not typical diagnostic test accuracy studies (62) as they did not report the sensitivity and specificity of tests (or positive and negative predictive values), nor did they explicitly refer to the use of a credible reference (gold) standard. Thus, some of the items in the EGAPP tool (66) and the STARD checklist (67) were not directly applicable. However, we judged both tools to be relevant for assessing the overall quality of the included studies. Further, through the use of these tools, weaknesses in conduct and/or reporting were highlighted, which potentially may be of help in improving future research in the field. Since there is overlap between some of the items/regions of the two tools, it may be worth considering using only one of them, or constructing a single instrument based on both.

One narrative review was included to address the advantages and limitations of tests for the detection of PIK3CA mutations. While systematic reviews and meta-analyses are considered the highest level of evidence, narrative reviews and expert opinion papers are generally considered to be low-level evidence (78). Normally we would not include this type of publications in a HTA, but as they appear to provide valuable information on the feasibility of tests, we chose to include them, and information from personal communication with Norwegian experts, to demonstrate the construct of the new framework we are piloting in this HTA.

The fact that the prevalence of PIK3CA gene mutations is higher in the ER+/HER2- subtype than in the other BC subtypes (11) is of importance for test accuracy, as higher prevalence typically is reflected by increased test sensitivity (68). In two of the larger studies (40, 41), all or a majority of patients belonged to the ER+/HER2- subtype, while in the smallest study (39), in which all patients were ER+, HER2 receptor status was not reported (authors were contacted for additional information but did not respond to our request). In addition, most participants in the included studies had progressed under endocrine treatment. The age of participants was not reported in two studies, and in one study both pre- and postmenopausal women

were included. We however believe that even though the included study populations do not exactly match our PICO, they are sufficiently representative for the BC group of interest for our review (i.e. patients with ER+/HER2- locally advanced or metastatic BC that have progressed under endocrine treatment). None of the studies reported the participants' progesterone hormone status, but the added value of this information for the results of this review is unclear.

Surprisingly, no studies comparing the accuracy of tests for the detection of PIK3CA mutations in tumour tissue specimens were identified, which both is a limitation with this review, and the research in this field. It is also problematic for this commission, and the use of this report, since tissue samples are used for all diagnostics at Norwegian hospital trusts, and liquid samples only for patient follow-up (46). Test accuracy is typically lower when liquid samples are used, as compared to tumour tissue (42), due to lower concentration of tumor-derived DNA (ctDNA) in plasma. We have not in our review been able to determine how big this difference is in terms of accuracy. Results from a systematic review and meta-analysis (69), which was not included in this review since it did not include a comparison between different tests, suggest high sensitivity and specificity (0.91 [0.58 to 0.99] and 0.98 [0.78 to 1.0] respectively), but also substantial heterogeneity (I^2 85% to 92%), for the detection of PIK3CA mutations in ctDNA samples as compared to in tissue samples from patients with metastatic BC. The five studies included in the analysis used various molecular tests (RT-PCR, Whole genome Sequencing, and NGS), had small sample sizes (N from 17 to 51), were relatively old (published between 2010 and 2014), and did not provide information on the molecular BC subtypes of participants. Another recent review (42), proposes a traffic-light system to guide what type of samples that safely can be used by different type of tests (RT-PCR, NGS, and Sanger) without risking false-negative results, e.g. was the use of RT-PCR for liquid samples highly recommended, while NGS analyses using liquid samples was suggested to result in increased risk of false negative results. For tissue samples it was the other way around i.e. NGS analysis was suggested to be more accurate, while RT-PCR using tissue samples would increase the risk of false negative results. For Sanger the use of liquid samples was discouraged, and tissue samples should be used with caution. It was however unclear what evidence this traffic-light system was based on. Since our review only included one study that compared ddPCR with NGS for the detection of PIK3CA mutations in plasma only, we cannot confirm or refute any of these propositions.

We did not search for studies of ELSI related to pharmacogenomic testing as we believe that ELSI due to its complexity (i.e. spanning factors like confidentiality issues, disclosure of genomic test results, ownership of data, panel testing for mutations with no available treatment, and the costs of scaling up of pharmacogenomic testing and targeted therapies (51), need to be addressed in a separate publication.

Nor did we search for studies concerned with patient preferences related to pharmacogenomic testing, or contacted patient consumers, which may be considered a

limitation with this report. The interested reader is referred to our previous publication (1), in which patient preferences were discussed.

As mentioned earlier, we did not find any studies that compared the value of the test-treatment combination based on the end-to-end studies and thus have no evidence concerning outcomes of importance to patients (e.g., overall survival, quality of life), or on how well different tests can predict the effectiveness of treatment (e.g., shrinking of the tumour, or slowing down the disease process).

***The included evidence**:*

Only three studies, which all provided inadequate evidence for test accuracy according to the EGAPP quality assessment, were included in this HTA. The quality of reporting was overall poor. Single studies, with relatively small sample sizes, provided evidence for each comparison. A sample size of at least 300 patients has been suggested as a rule of thumb for studies aimed at assessing test accuracy (70), but only one of the included studies fulfilled the sample size criteria.

The quality of the included evidence for the accuracy of tests from the three studies was overall inadequate. None of the included studies reported accuracy statistics (such as sensitivity and specificity), which typically is reported in diagnostic accuracy studies (62), but reported concordance, even if there are other, possible more robust, statistical methods that can be used when no credible reference standard is available (63).

None of the included studies described a hypothesis or a sample size calculation, which is in line with what has been reported in the literature (62) i.e. that study hypotheses are rarely pre-defined and sample size calculations are usually not performed in diagnostic accuracy studies, which according to the authors, and I quote: "*may jeopardize scientific rigor and can lead to over-interpretation or "spin" of study findings*". The authors (62) further suggest that minimally acceptable criteria for test performance should form the basis for both the formulation of the hypothesis and the sample size calculations in all diagnostic accuracy studies.

Test accuracy may vary depending on many factors, e.g. included population, the spectrum of disease, test interpreters, and the results of prior testing (71). The included studies however, provided patchy information on patient selection and/or the patient characteristics, on blinding of test assessors (only one study (40) explicitly stated that the testing was blinded), and on where in the clinical pathway the tests were performed.

In addition, none of the studies verified all cases, and typically provided little or no information on how missing or indeterminate samples were handled, while ignoring indeterminate test results may produce biased estimates of accuracy (72). The quality of reporting was poor.

As mentioned earlier, none of the three included studies reported on the sensitivity and specificity of tests, but instead reported concordance (agreement between tests) which typically is used to establish the validity of a new diagnostic technique or to demonstrate the equivalence across analytical techniques (73). One of the included studies (41) highlighted the lack of an adequate reference standard test for the detection of PIK3CA mutations in plasma, which probably was the reason why all three studies reported concordance. However, concordance according to Chikere and colleagues (63), and I cite, "*is best used as an exploratory tool alongside other methods as it lacks robustness to assess the diagnostic ability of a test*". Chikere et al. describe in their review (63) four main groups of statistical methods that can be used for the evaluation of medical test(s) when there is no available reference (gold) standard. The authors further provide a flow-diagram to assist researchers in choosing appropriate statistical methods in different scenarios, which potentially may be of help in future diagnostic test accuracy studies (63).

Economic evaluation

Several testing strategies for the detection of PIK3CA gene mutations in locally advanced or metastatic BC exist, however there is still a lack of evidence on the comparative economic implications of using these strategies in Norwegian clinical practice.

We assessed the costs associated with the most relevant diagnostic methods for the detection of PIK3CA mutations in advanced BC in Norwegian clinical settings, i.e., RT-PCR and NGS. The analyses were based on the data received from the Norwegian University Hospitals, the laboratories of molecular biology.

Liquid or tumour tissue samples from patients with MBC can be used for the detection of PIK3CA mutations. However, the test accuracy is reported to be typically lower when liquid samples are used, as compared to tumour tissue (42). It is therefore recommended that in case of negative results using liquid samples (both for PCR and NGS analyses), a tissue biopsy should be also analysed to avoid false-negative results (45). Thus, we expect that the costs associated with PIK3CA mutations testing in liquid samples are higher than testing in tissue samples.

According to our experts, tissue samples are used for all diagnostics at Norwegian hospital trusts, while liquid samples mainly are used for patient follow-up. In addition, due to uncertainty associated with data on resource use and data related to diagnostic accuracy of the methods using liquid samples for PIK3CA testing, our micro-cost analysis is limited to the cost associated with RT-PCR or NGS testing using the tissue samples.

Using RT-PCR for the detection of gene mutation has some advantages. It has been reported that if a known mutation is being followed through treatment (e.g., PIK3CA mutation), a PCR-focused test is usually more cost-effective (42). However, the technology also has intrinsic limitations, including its limited multiplexing capability. In addition, the most commercially available PCR kits for PIK3CA mutational analysis do not provide quantitative information about mutant allele frequency (42). The results of our cost analysis showed that using PCR for the detection of several mutations simultaneously will increase the reagent costs, and consequently, the total costs compared to testing one biomarker with PCR.

Further, our assessment showed that the costs with the NGS approach are likely to be higher than the other diagnostic methods if it is used for one patient only, due to the higher reagent costs of the NGS test. However, as NGS technology allows massively parallel sequencing and testing samples from several patients at once, the cost associated with NGS testing will significantly be decreased when parallel tests are performed on several biomarkers from multiple patients. In addition, the need for tissue preservation and the burden and comorbidity of repeat biopsies is likely to decrease. The results showed that the cost associated with using NGS was around NOK 16,000 per sample. However, if several samples and biomarkers including PIK3CA mutations in MBC patients are tested at the same time with NGS, the costs will be reduced to approximately NOK 2,000 per patient.

Assuming about 140 patients annually are eligible for testing to detect of PIK3CA mutations in patients with advanced BC, the costs were estimated to be approximately NOK 322,000. The costs did not include overhead costs, capital and other infrastructure costs.

Based on the data from the Norwegian pathology departments, the capital and infrastructure costs (including overhead costs) are higher for NGS than the other diagnostic methods. It has been estimated that the costs related to NGS equipment and supplies investments is approximately NOK 3-4 million. Further, the costs associated to maintenance the NGS instruments is more expensive (around NOK 30,000- 80,000) than the other methods. According to the experts, the validation process for any of the techniques can be challenging. When there are not enough positive cases, the validation process at local hospital is dependent on positive cases/controls from other countries/companies. Although using NGS technology requires highly trained personnel and the infrastructure and maintenance costs associated with this technology still are high, the optimisation of the laboratory workflow and volumes allows for more efficient use of the resources.

A comparison between the estimated costs based on the micro-costing method and the costs estimated based on the new reimbursement scheme (59) indicated that the reimbursement rates are generally insufficient to cover the costs of running PCR. For NGS testing, the analyses showed that the reimbursement rates can cover the costs of

running the small NGS assay. In addition, the expanded gene panel used for the identification of the relevant cancer patients mainly in research has recently received new reimbursement codes which can be used for the larger NGS panels.

Multiple labs also reported differences in reimbursement for internal versus external testing (inpatient versus outpatient) which may led to apparently unintended consequences, such as clinicians triaging groups of patients to prioritize for external testing when assays were not available in-house, or electing to perform diagnostics in an outpatient rather than inpatient setting if possible (74).

Although, we have tried to conduct our analysis based on the best available data, limited access to the cost data comparing different diagnostic methods and data related to the diagnostic accuracy of the methods using liquid samples, were the most important limitation of this economic analysis. We contacted all Norwegian regional health authorities. We received data from three hospitals based on the relative different approaches. Due to data consistency, our analyses are based on the data received from Stavanger University Hospital. Furthermore, we have not considered the additional re-biopsy costs in our analyses.

There is some uncertainty regarding the number of eligible patients for PIK3CA mutation testing in Norway, due to the lack of information on the number of patients with locally advanced or metastatic BC, as well as the number of patients who actually received and progressed during endocrine therapy, in national registries. Our estimates however are based on Norwegian experts' opinions.

Our systematic search did not identify any studies that compared the value of the test-drug combinations. In line with the commission, in this HTA, we have estimated the costs associated with the diagnostic methods for the detection of PIK3CA mutations that are relevant in a Norwegian context. In parallel, based on the clinical data presented in the SOLAR-1 publications (24, 75), we will develop a health economic decision model to assess the value of the test-treatment combination.

Challenges

Recently, the number of economic evaluations on precision medicine has noticeably increased. However, an important number of these analyses focus purely on the assessment of the actual therapeutic treatment, failing to include the impact that the actual tests have on the overall economic value of the test–drug combination (76). Even, when both testing and therapeutic decisions are taken into consideration, the weight of the accuracy of the companion tests on the overall results is rarely explored (76).

In principle, the healthcare system benefits from the availability of companion diagnostics that accurately identify responders, reduce the number needed to treat, and thereby improve the efficient use of resource. Therefore, the consequences of

implementation of test-and-treat interventions and system integration challenges should be considered by the reimbursement authorities. In addition to that a good biomarker test should measure the biomarker with a high degree of accuracy and demonstrate analytical validity, the biomarker test should optimally demonstrate clinical utility, meaning that it improves patient's outcomes compared to a no-testing approach. This requires establishing the relationship between the test results and the consecutive treatment, and outcomes. Generating evidence to support the economic case of precision medicine in practice, however, can be a challenge. Manufacturers, analysts, and funders of research may improve their research and development activities by considering the evidenced required by later-stage decision-makers at an earlier time period in the process of evidence generation.

The basic principles of cost-effectiveness should be applied to biomarkers. However, cost-effectiveness estimates for recent pharmaceutical-diagnostic combinations have been variable among major HTA organisations, suggesting that methods for incorporating test information into economic evaluations are inconsistent. Key issues include gaps in the evidence supporting clinical utility and cost effectiveness of diagnostics (77).

Limited information on treatment patterns and on the costs and outcomes related to the use of different diagnostic methods, is a major problem with the published economic evaluation of precision medicine interventions, especially the lack of data about false-positive and false-negative test results is problematic (77). A higher specificity rate of a diagnostic method will help to reduce the potential treatment of 'false positives' and consequently engaging in high spending for a proportion of the patient population for which the targeted therapy would not be effective or for which it could even produce some harm. San Miguel (76) have shown that in the field of targeted therapy in oncology, test accuracy becomes even more crucial given high effectiveness but also high prices of some of these therapies. The specificity of a diagnostic method becomes even more important if very small population subsets are to be identified using the biomarker. San Miguel (76) concluded that the importance of test specificity is twofold; for the patient it is crucial to receive the correct targeted treatment; for the society the use of (often expensive) targeted treatment in patients that do not benefit from it because the marker was not tested accurately, is a waste of money.

Furthermore, the health economic evaluation can even become more complex if different tests are combined or sequentially used. This potential complexity can be handled by explicitly showing how these tests are going to be used in practice and then working with the combined sensitivities and specificities of the tests (53). Moreover, the tests available and the test sequence employed in the clinical study, may differ in their ability to accurately select patients who are most likely to benefit from targeted therapy.

In addition to the issues related to properly assessing the health economic consequences of test-drug mentioned above, there are regulatory barriers to the development and adoption of precision medicine. First, the regulation of marketing approval is insufficiently harmonized. It varies across countries and is different for drugs and diagnostic tests. In the United States, marketing approval for drugs and diagnostics is done by the FDA. The joint approval process performed by a single agency ensures scientific knowledge-sharing and provides an effective way to approve precision medicines. However, in Europe, no single European agency regulates both medicines and tests. The European Medicines Agency (EMA) regulates the marketing approval of drugs, whereas it is each European Union (EU) member state's Notified Body that monitors the performance standards of diagnostic test (78).

Furthermore, the requirements for marketing approval of tests are still relatively lenient. In Europe, the test manufacturer is currently required to demonstrate the clinical validity (predictive capability) but not the clinical utility (effect on clinical outcomes) of the test. Another important challenge related to assess the cost-effectiveness of diagnostic methods is that, currently, the manufacturer of a new test does not need to demonstrate its effectiveness if a similar test already exists. Moreover, laboratory developed tests, that is, tests performed within a single laboratory or hospital (not commercialized) do not require a full regulatory review (78). Therefore, there is a lack of standardized evidence of the performance of biomarker tests in terms of their impact on health outcomes. This results in uncertainty for health authorities who make decisions on pricing and reimbursement based on the value of treatment produced by the biomarker test.

However, in 2017 the EU parliament and council agreed on a new set of regulations on in vitro diagnostics (79). Based on these new regulations, companion diagnostics will need to meet stricter performance requirements, including clinical evidence and there will be a link between the assessment of a diagnostics by a notified body and the corresponding medical product by a medicine regulatory authority. It means that if a companion diagnostic is necessary to identify whether a patient is likely to benefit from a corresponding medical product, the evidence regarding its impact on patient outcome, i.e., clinical utility, will be carefully considered by the medicine authorities, in determining the benefit/risk of the medical product. Consequently, these evidence on clinical utility of companion diagnostics can be used in evaluation of health economic consequences of the test and the consecutive treatment. The regulations will become full effective in May 2022 (79).

Conclusion

Results of this HTA is based on inadequate evidence on the accuracy of tests (various PCR assays and NGS-panels) for detection of PIK3CA mutations from three studies. Un-pooled results (5 tests and 3 comparisons) from single studies, with relatively small sample sizes, suggest good concordance between tests (Cohen's κ: 0.80 to 0.86). All studies used plasma samples only.

We did not identify any systematic reviews, or meta-analyses evaluating the accuracy of tests for the detection of PIK3CA mutations. None of the included studies reported on clinical utility. No evidence for the test accuracy of Sanger sequencing or Liquid chip methodologies were found.

In our cost analysis, we included the costs for the tests of most relevance for a Norwegian context. Our results showed that for the detection of PIK3CA in isolation, the costs for testing using PCR is less than NGS-panel testing, however, using PCR for the detection of additional relevant mutations, will increase total cost. At present, the capital and infrastructure as well as maintenance costs are higher for NGS than the other diagnostic methods. Assuming that about 140 patients with metastatic BC annually are eligible for PIK3CA mutation testing annually, the costs were estimated to be ca NOK 322,000.

The new reimbursement rates are insufficient to cover the costs of running PCR. The reimbursement rates for NGS testing can cover the costs of running the small NGS assays. In addition, new reimbursement rates have recently been defined for the expanded gene panels that mainly are used for the experimental treatment.

All tests have both advantages and limitations, but due to incomplete information an adequate comparison was unfeasible. However, the choice of a suitable test for detecting PIK3CA mutations depends on accessibility of testing modalities, economic considerations, sample type and risk of false negatives, and the TAT.

Future research should focus on conducting larger cohort studies with well-defined patient populations, that follows the patients from testing (or no testing), through treatment and final outcomes. Further, robust and replicable methods, as well as a reporting standard checklist, should be used for increased clarity.

References

1. Flodgren G HV. Tests for detection of ROS1 gene alterations in people with non-small cell lung cancer (NSCLC). A health technology assessment Oslo, Norway: Norwegian Insitutute of Public Health;; 2021.
2. Awad K DM, Cree I, Challoner B, Ghosh S, Thurston D,. The precision medicine approach to cancer therapy: part 1 — solid tumours. . The Pharmaceutical Journal 2019.
3. Robinson DR, Wu YM, Lonigro RJ, Vats P, Cobain E, Everett J, et al. Integrative clinical genomics of metastatic cancer. Nature. 2017;548(7667):297-303.
4. Califf RM. Biomarker definitions and their applications. Exp Biol Med (Maywood). 2018;243(3):213-21.
5. Tsimberidou A-M, Hong DS, Wheler JJ, Falchook GS, Janku F, Naing A, et al. Long-term overall survival and prognostic score predicting survival: the IMPACT study in precision medicine. Journal of Hematology & Oncology. 2019;12(1):145.
6. Keeling P, Clark J, Finucane S. Challenges in the clinical implementation of precision medicine companion diagnostics. Expert Review of Molecular Diagnostics. 2020;20(6):593-9.
7. Burke W, Atkins D, Gwinn M, Guttmacher A, Haddow J, Lau J, et al. Genetic test evaluation: information needs of clinicians, policy makers, and the public. Am J Epidemiol. 2002;156(4):311-8.
8. Phillips KA, Deverka PA, Hooker GW, Douglas MP. Genetic Test Availability And Spending: Where Are We Now? Where Are We Going? Health Aff (Millwood). 2018;37(5):710-6.
9. IQN Path E, EFPIA,. Improving cancer care through broader access to quality biomarker testing EFPIA; 2021 [
10. Feng Y, Spezia M, Huang S, Yuan C, Zeng Z, Zhang L, et al. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes Dis. 2018;5(2):77-106.
11. American Cancer Society (ACS). Breast Cancer Facts & Figures 2019-2020. Atlanta: ; 2021.
12. Helsedirektoratet. Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av pasienter med brystkreft. Oslo, Norway; 2021.
13. Escala-Garcia M, Canisius S, Keeman R, Beesley J, Anton-Culver H, Arndt V, et al. Germline variants and breast cancer survival in patients with distant metastases at primary breast cancer diagnosis. Sci. 2021;11(1).
14. Harbeck N, Gnant M. Breast cancer. The Lancet. 2017;389(10074):1134-50.
15. MedlinePlus [Internet]. PIK3CA gene phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha Bethesda (MD): : National Library of Medicine (US); ; 2022 [[updated Jun 24; cited 2020 Jul 1]. Available from: https://medlineplus.gov/.].
16. Zhang Z, Richmond A. The Role of PI3K Inhibition in the Treatment of Breast Cancer, Alone or Combined With Immune Checkpoint Inhibitors. Front Mol Biosci. 2021;8:648663.

17. Martínez-Sáez O, Chic N, Pascual T, Adamo B, Vidal M, González-Farré B, et al. Frequency and spectrum of PIK3CA somatic mutations in breast cancer. Breast Cancer Res. 2020;22(1):45.
18. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, McMichael JF, et al. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61-70.
19. Zardavas D, Marvelde Lt, Milne RL, Fumagalli D, Fountzilas G, Kotoula V, et al. Tumor PIK3CA Genotype and Prognosis in Early-Stage Breast Cancer: A Pooled Analysis of Individual Patient Data. J Clin Oncol. 2018;36(10):981-90.
20. Anderson EJ, Mollon LE, Dean JL, Warholak TL, Aizer A, Platt EA, et al. A Systematic Review of the Prevalence and Diagnostic Workup of PIK3CA Mutations in HR+/HER2- Metastatic Breast Cancer. Int J Breast Cancer. 2020;2020:3759179-.
21. Hartkopf AD, Grischke E-M, Brucker SY. Endocrine-Resistant Breast Cancer: Mechanisms and Treatment. Breast Care (Basel). 2020;15(4):347-54.
22. Norvegian Medicines Agency (NoMA). Palbociklib i kombinasjon med fulvestrant til behandling av HR-positiv, HER2-negativ metastatisk brystkreft som har progrediert etter endokrin behandling. Oslo, Norway; 2019.
23. Cardoso F, Paluch-Shimon S, Senkus E, Curigliano G, Aapro MS, André F, et al. 5th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 5). Ann Oncol. 2020;31(12):1623-49.
24. André F, Ciruelos E, Rubovszky G, Campone M, Loibl S, Rugo HS, et al. Alpelisib for PIK3CA-Mutated, Hormone Receptor–Positive Advanced Breast Cancer. New England Journal of Medicine. 2019;380(20):1929-40.
25. Ardini E, Siena S. Entrectinib approval by EMA reinforces options for ROS1 and tumour agnostic NTRK targeted cancer therapies. ESMO Open. 2020;5(5):e000867.
26. US Food and drug adminisatration (FDA). FDA approves entrectinib for NTRK solid tumors and ROS-1 NSCLC 2019 [
27. Wong ML, Medrano JF. Real-time PCR for mRNA quantitation. BioTechniques. 2005;39(1):75-85.
28. Behjati S, Tarpey PS. What is next generation sequencing? Arch Dis Child Educ Pract Ed. 2013;98(6):236-8.
29. Ming Z, Jiang D, Hu Q, Li X, Huang J, Xu Y, et al. Diagnostic application of PIK3CA mutation analysis in Chinese esophageal cancer patients. Diagnostic Pathology. 2014;9(1):153.
30. FDA. List of Cleared or Approved Companion Diagnostic Devices (In Vitro and Imaging Tools) USA: US Food and drug administration; 2021 [
31. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genet Med. 2009;11(1):3-14.
32. Veenstra DL, Piper M, Haddow JE, Pauker SG, Klein R, Richards CS, et al. Improving the efficiency and relevance of evidence-based recommendations in the era of whole-genome sequencing: an EGAPP methods update. Genetics in Medicine. 2013;15(1):14-24.
33. Pitini E, D'Andrea E, De Vito C, Rosso A, Unim B, Marzuillo C, et al. A proposal of a new evaluation framework towards implementation of genetic tests. PloS one. 2019;14(8):e0219755-e.
34. Huang X, Lin J, Demner-Fushman D. Evaluation of PICO as a knowledge representation for clinical questions. Annual Symposium proceedings AMIA Symposium. 2006;2006:359-63.
35. the EndNote team. EndNote. EndNote X9 ed. Philadelphia, PA: Clarivate; 2013.
36. Ouzzani M HH, Fedorowicz Z, and Elmagarmid A. Rayyan — a web and mobile app for systematic reviews. Systematic Reviews. 2016;5(210).
37. Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. Plast Reconstr Surg. 2011;128(1):305-10.

38. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Bmj. 2015;351:h5527.
39. Nteliopoulos G, Page K, Hills A, Howarth K, Emmett W, Green E, et al. Comparison of two targeted ultra-deep sequencing technologies for analysis of plasma circulating tumour DNA in endocrine-therapy-resistant breast cancer patients. Breast Cancer Res Treat. 2021;188(2):465-76.
40. O'Leary B, Hrebien S, Beaney M, Fribbens C, Garcia-Murillas I, Jiang J, et al. Comparison of BEAMing and Droplet Digital PCR for Circulating Tumor DNA Analysis. Clin Chem. 2019;65(11):1405-13.
41. Zivanovic Bujak A, Weng C-F, Silva MJ, Yeung M, Lo L, Ftouni S, et al. Circulating tumour DNA in metastatic breast cancer to guide clinical trial enrolment and precision oncology: A cohort study. PLOS Medicine. 2020;17(10):e1003363.
42. Fusco N, Malapelle U, Fassan M, Marchiò C, Buglioni S, Zupo S, et al. PIK3CA Mutations as a Molecular Target for Hormone Receptor-Positive, HER2-Negative Metastatic Breast Cancer. Front Oncol. 2021;11:644737.
43. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
44. Markou A, Tzanikou E, Ladas I, Makrigiorgos GM, Lianidou E. Nuclease-Assisted Minor Allele Enrichment Using Overlapping Probes-Assisted Amplification-Refractory Mutation System: An Approach for the Improvement of Amplification-Refractory Mutation System-Polymerase Chain Reaction Specificity in Liquid Biopsies. Analytical Chemistry. 2019;91(20):13105-11.
45. Toppmeyer DL, Press MF. Testing considerations for phosphatidylinositol-3-kinase catalytic subunit alpha as an emerging biomarker in advanced breast cancer. Cancer medicine. 2020;9(18):6463-72.
46. Janssen E. Personal communication regarding using molecular tests in a Norwegian context. Stavanger University Hospital; 2021-2022.
47. Oslo universitetssykehus. NGS diagnostikk innen kreft (somatiske forendringer). Oslo, Norway; 2020.
48. Ministry of Health and Care Services. A white paper on priority setting in the Norwegian health care sector. 2016. Contract No.: Meld. St. 34 (2015–2016).
49. Shabaruddin FH, Fleeman ND, Payne K. Economic evaluations of personalized medicine: existing challenges and current developments. Pharmgenomics Pers Med. 2015;8:115-26.
50. Hatz MH, Schremser K, Rogowski WH. Is individualized medicine more cost-effective? A systematic review. Pharmacoeconomics. 2014;32(5):443-55.
51. Kasztura M, Richard A, Bempong NE, Loncar D, Flahault A. Cost-effectiveness of precision medicine: a scoping review. Int J Public Health. 2019;64(9):1261-71.
52. National Institute for Health and Clinical Excellence. Diagnostic Assessment Programme Manual. 2011 2011.
53. Annemans L, Redekop K, Payne K. Current methodological issues in the economic assessment of personalized medicine. Value Health. 2013;16(6 Suppl):S20-6.
54. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes: Oxford University Press; 2015.
55. Kievits T, Niese D, L.T. H, P. C, Gledic S, Roediger A, et al. Personalized Medicine:Status qua and Challenges. The European Association for Bioindustries; 2021.
56. Cancer Registry of Norway. Cancer in Norway. Cancer Registry of Norway. 2021.
57. Brystkreftforeningen. Om brystkreft. 2020.
58. Hamidi V FG, Meneses-Echavez JF, Bidonde J,. . Tests for the detection of NTRK gene fusions in patients with locally advanced or metastatic solid tumours. Oslo, Norway: Norwegian Institute of Public Health,; 2022.
59. Helse- og omsorgsdepartementet. Forskrift om godtgjørelse for å yte poliklinisk helsehjelp i spesialisthelsetjenesten (poliklinikkforskriften). 2022.

60. Helfo. Ny refusjonsordning for poliklinisk patologi 2022 [Available from: https://www.helfo.no/regelverk-og-takster/overordnet-regelverk/regelverk-og-refusjon-for-sjukehus-og-poliklinikk/ny-refusjonsordning-for-poliklinisk-patologi.
61. Norwegian Medicine Agency. Retningslinjer for dokumentasjonsgrunnlag for hurtig metodevurdering av legemidler. Oslo: Norwegian Medicine Agency; 2021.
62. Korevaar DA, Gopalakrishna G, Cohen JF, Bossuyt PM. Targeted test evaluation: a framework for designing diagnostic accuracy studies with clear study hypotheses. Diagnostic and Prognostic Research. 2019;3(1):22.
63. Umemneku Chikere CM, Wilson K, Graziadio S, Vale L, Allen AJ. Diagnostic test evaluation methodology: A systematic review of methods employed to evaluate diagnostic tests in the absence of gold standard - An update. PLoS One. 2019;14(10):e0223832.
64. Mosele F, Remon J, Mateo J, Westphalen CB, Barlesi F, Lolkema MP, et al. Recommendations for the use of next-generation sequencing (NGS) for patients with metastatic cancers: a report from the ESMO Precision Medicine Working Group. Ann Oncol. 2020;31(11):1491-505.
65. Hamidi V FG, Meneses-Echavez JF, Bidonde J,. Tests for the detection of NTRK gene fusions in patients with locally advanced or metastatic solid tumours. A Health Technology Assessment. Oslo, Norway: Norwegian Institute of Public Health; 2022.
66. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genetics in medicine : official journal of the American College of Medical Genetics. 2009;11(1):3-14.
67. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ : British Medical Journal. 2015;351:h5527.
68. Leeflang MMG, Rutjes AWS, Reitsma JB, Hooft L, Bossuyt PMM. Variation of a test's sensitivity and specificity with disease prevalence. CMAJ. 2013;185(11):E537-E44.
69. Zhou Y, Wang C, Zhu H, Lin Y, Pan B, Zhang X, et al. Diagnostic Accuracy of PIK3CA Mutation Detection by Circulating Free DNA in Breast Cancer: A Meta-Analysis of Diagnostic Test Accuracy. PLoS One. 2016;11(6):e0158143-e.
70. Bujang MA, Adnan TH. Requirements for Minimum Sample Size for Sensitivity and Specificity Analysis. J Clin Diagn Res. 2016;10(10):YE01-YE6.
71. Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM, Cochrane Diagnostic Test Accuracy Working G. Systematic reviews of diagnostic test accuracy. Annals of internal medicine. 2008;149(12):889-97.
72. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. 2016;6(11):e012799.
73. Kwiecien R, Kopp-Schneider A, Blettner M. Concordance analysis: part 16 of a series on evaluation of scientific publications. Dtsch Arztebl Int. 2011;108(30):515-21.
74. DNV GL.-BigMed. Implementing NGS-based diagnostics in cancer care: Technical and organizational factors in the Nordics. 2021.
75. Andre F, Ciruelos EM, Juric D, Loibl S, Campone M, Mayer IA, et al. Alpelisib plus fulvestrant for PIK3CA-mutated, hormone receptor-positive, human epidermal growth factor receptor-2-negative advanced breast cancer: final overall survival results from SOLAR-1. Ann Oncol. 2021;32(2):208-17.
76. San Miguel L, Hulstaert F. The importance of test accuracy in economic evaluations of companion diagnostics. J Comp Eff Res. 2015;4(6):569-77.
77. Faulkner E, Annemans L, Garrison L, Helfand M, Holtorf AP, Hornberger J, et al. Challenges in the development and reimbursement of personalized medicine-payer and manufacturer perspectives and implications for health economics and outcomes research: a report of the ISPOR personalized medicine special interest group. Value Health. 2012;15(8):1162-71.

78. Meadows NA, Morrison A, Brindley DA, Schuh A, Barker RW. An evaluation of regulatory and commercial barriers to stratified medicine development and adoption. Pharmacogenomics J. 2015;15(1):6-12.
79. The European parliament and of the council of the European Union. Regulation (EU) 2017/746 of the European parliament and of the council of 5 April 2017 on in vitro diagnostic medical devices
2017.
80. Alvarez-Garcia V, Bartos C, Keraite I, Trivedi U, Brennan PM, Kersaudy-Kerhoas M, et al. A simple and robust real-time qPCR method for the detection of PIK3CA mutations. Sci. 2018;8(1):4290.
81. Andre F, Bachelot T, Commo F, Campone M, Arnedos M, Dieras V, et al. Comparative genomic hybridisation array and DNA sequencing to direct treatment of metastatic breast cancer: a multicentre, prospective trial (SAFIR01/UNICANCER). Lancet Oncol. 2014;15(3):267-74.
82. Ang D, O'Gara R, Schilling A, Beadling C, Warrick A, Troxell ML, et al. Novel method for PIK3CA mutation analysis: locked nucleic acid--PCR sequencing. Journal of Molecular Diagnostics. 2013;15(3):312-8.
83. Arnedos M, Scott V, Job B, De La Cruz J, Commo F, Mathieu MC, et al. Array CGH and PIK3CA/AKT1 mutations to drive patients to specific targeted agents: a clinical experience in 108 patients with metastatic breast cancer. Eur J Cancer. 2012;48(15):2293-9.
84. Arsenic R, Treue D, Lehmann A, Hummel M, Dietel M, Denkert C, et al. Comparison of targeted next-generation sequencing and Sanger sequencing for the detection of PIK3CA mutations in breast cancer. BMC Clin Pathol. 2015;15:20.
85. Beaver JA, Jelovac D, Balukrishna S, Cochran R, Croessmann S, Zabransky DJ, et al. Detection of cancer DNA in plasma of patients with early-stage breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2014;20(10):2643-50.
86. Bergholtz H, Kumar S, Wärnberg F, Lüders T, Kristensen V, Sørlie T. Comparable cancer-relevant mutation profiles in synchronous ductal carcinoma in situ and invasive breast cancer. Cancer Rep (Hoboken). 2020;3(3):e1248.
87. Bertucci F, Finetti P, Guille A, Adelaide J, Garnier S, Carbuccia N, et al. Comparative genomic analysis of primary tumors and metastases in breast cancer. Oncotarget. 2016;7(19):27208-19.
88. Bertucci F, Gonçalves A, Guille A, Adelaïde J, Garnier S, Carbuccia N, et al. Prospective high-throughput genome profiling of advanced cancers: results of the PERMED-01 clinical trial. Genome Medicine. 2021;13(1):87.
89. Board RE, Wardley AM, Dixon JM, Armstrong AC, Howell S, Renshaw L, et al. Detection of PIK3CA mutations in circulating free DNA in patients with breast cancer. Breast Cancer Research and Treatment. 2010;120(2):461-7.
90. Chin YM, Takahashi Y, Chan HT, Otaki M, Fujishima M, Shibayama T, et al. Ultradeep targeted sequencing of circulating tumor DNA in plasma of early and advanced breast cancer. Cancer Sci. 2021;112(1):454-64.
91. Dawson S-J, Tsui DWY, Murtaza M, Biggs H, Rueda OM, Chin S-F, et al. Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. New England Journal of Medicine. 2013;368(13):1199-209.
92. De Angelis C, Bruzzone M, Caparica R, Crucitta S, Ceppi M, Fontana A, et al. 101P Accuracy of detecting PIK3CA mutations on circulating tumour (ct) DNA as compared to tumour tissue (TT) in breast cancer (BC) patients (pts): Results from a systematic review and meta-analysis. Annals of Oncology. 2020;31:S281-S2.
93. Dirican E, Kaya Z, Gullu G, Peker I, Ozmen T, Gulluoglu BM, et al. Detection of PIK3CA gene mutations with HRM analysis and association with IGFBP-5 expression levels in breast cancer. Asian Pac J Cancer Prev. 2014;15(21):9327-33.
94. Dumbrava EE, Call SG, Huang HJ, Stuckett AL, Madwani K, Adat A, et al. PIK3CA mutations in plasma circulating tumor DNA predict survival and treatment outcomes in patients with advanced cancers. ESMO Open. 2021;6(5):100230.

95. Gerratana L, Zhang Q, Shah AN, Davis AA, Zhang Y, Wehbe F, et al. Performance of a novel Next Generation Sequencing circulating tumor DNA (ctDNA) platform for the evaluation of samples from patients with metastatic breast cancer (MBC). Critical Reviews in Oncology/Hematology. 2020;145(102856).
96. Gerratana L, Davis AA, Velimirovic M, D'Amico P, Shah AN, Clifton K, et al. Uncovering the differential impact of ESR1 and PIK3CA codon variants on the clinical phenotype of metastatic breast cancer (MBC) through circulating tumor DNA (ctDNA) next-generation sequencing (NGS). J Clin Oncol. 2021;39(15_suppl):1033-.
97. Harle A, Lion M, Lozano N, Husson M, Harter V, Genin P, et al. Analysis of PIK3CA exon 9 and 20 mutations in breast cancers using PCR-HRM and PCR-ARMS: correlation with clinicopathological criteria. Oncol Rep. 2013;29(3):1043-52.
98. Higgins MJ, Jelovac D, Barnathan E, Blair B, Slater S, Powers P, et al. Detection of tumor PIK3CA status in metastatic breast cancer using peripheral blood. Clinical Cancer Research. 2012;18(12):3462-9.
99. Hirsch B, Endris V, Lassmann S, Weichert W, Pfarr N, Schirmacher P, et al. Multicenter validation of cancer gene panel-based next-generation sequencing for translational research and molecular diagnostics. Virchows Archiv. 2018;472(4):557-65.
100. Ibrahim EM, Kazkaz GA, Al-Mansour MM, Al-Foheidi ME. The predictive and prognostic role of phosphatase phosphoinositol-3 (PI3) kinase (PIK3CA) mutation in HER2-positive breast cancer receiving HER2-targeted therapy: a meta-analysis. Breast Cancer Research and Treatment. 2015;152(3):463-76.
101. Kaur P, Porras TB, Ring A, Carpten JD, Lang JE. Comparison of TCGA and GENIE genomic datasets for the detection of clinically actionable alterations in breast cancer. Sci. 2019;9(1):1482.
102. Lambert A, Salleron J, Lion M, Rouyer M, Lozano N, Leroux A, et al. Comparison of Three Real-Time PCR Assays for the Detection of PIK3CA Somatic Mutations in Formalin-Fixed Paraffin Embedded Tissues of Patients with Breast Carcinomas. Pathol Oncol Res. 2019;25(3):1117-23.
103. Liu YR, Jiang YZ, Zuo WJ, Yu KD, Shao ZM. PIK3CA mutations define favorable prognostic biomarkers in operable breast cancer: a systematic review and meta-analysis. OncoTargets and therapy. 2014;7:543-52.
104. Ney JT, Froehner S, Roesler A, Buettner R, Merkelbach-Bruse S. High-Resolution Melting Analysis as a Sensitive Prescreening Diagnostic Tool to Detect KRAS, BRAF, PIK3CA, and AKT1 Mutations in Formalin-Fixed, Paraffin-Embedded Tissues. Arch Pathol Lab Med. 2012;136(9):983-92.
105. Ou SI, Nagasaka M, Zhu VW. Liquid Biopsy to Identify Actionable Genomic Alterations. Am. 2018;38:978-97.
106. Pang B, Cheng S, Sun S-P, An C, Liu Z-Y, Feng X, et al. Prognostic role of PIK3CA mutations and their association with hormone receptor expression in breast cancer: a meta-analysis. Sci. 2014;4(1):6255.
107. Raphael MJ, Robinson A, Booth CM, O'Donnell J, Palmer M, Eisenhauer E, et al. The Value of Progression-Free Survival as a Treatment End Point Among Patients With Advanced Cancer: A Systematic Review and Qualitative Assessment of the Literature. JAMA Oncol. 2019;5(12):1779-89.
108. Rothé F, Laes JF, Lambrechts D, Smeets D, Vincent D, Maetens M, et al. Plasma circulating tumor DNA as an alternative to metastatic biopsies for mutational analysis in breast cancer. Ann Oncol. 2014;25(10):1959-65.
109. Sobhani N, Roviello G, Corona SP, Scaltriti M, Ianza A, Bortul M, et al. The prognostic value of PI3K mutational status in breast cancer: A meta-analysis. J Cell Biochem. 2018;119(6):4287-92.
110. Shimoi T, Hamada A, Yamagishi M, Hirai M, Yoshida M, Nishikawa T, et al. PIK3CA mutation profiling in patients with breast cancer, using a highly sensitive detection system. Cancer Sci. 2018;109(8):2558-66.

111. Song J, Zhang J, Lv F, Cheng Y, Wang B, Feng L, et al. Multiplex detection of DNA mutations by the fluorescence fingerprint spectrum technique. Angew Chem Int Ed Engl. 2013;52(49):13020-3.
112. Tay TKY, Tan PH. Liquid Biopsy in Breast Cancer: A Focused Review. Arch Pathol Lab Med. 2020;11:11.
113. Yang X, Zhang K, Zhang C, Peng R, Sun C. Accuracy of analysis of cfDNA for detection of single nucleotide variants and copy number variants in breast cancer. BMC Cancer. 2019;19(1):465.
114. Zhou Y, Wang C, Zhu H, Lin Y, Pan B, Zhang X, et al. Diagnostic Accuracy of PIK3CA Mutation Detection by Circulating Free DNA in Breast Cancer: A Meta-Analysis of Diagnostic Test Accuracy. PLoS ONE [Electronic Resource]. 2016;11(6):e0158143.

Appendices

Appendix 1 Glossary

Advanced cancer	is cancer has grown outside the body part it started in but has not yet spread to other body parts
Alpelisib	is a targeted therapy medicine, and a PI3K inhibitor that blocks the PI3K pathway
Analytical validity	or technical performance, is a test's ability to accurately and reliably *measure* a biomarker of interest (sensitivity, specificity, assay robustness, and quality control). According to EGAPP definitions https://www.cdc.gov/genomics/gtesting/egapp/recommend/method.htm
Breast carcinomas	a condition in which abnormal cells are found in the tissues of the breast
Gene alteration	a somatic gene alteration
Cell-free DNA (cfDNA)	is DNA that is freely circulating in the bloodstream that is not necessarily of tumor origin
Chemotherapy	a drug treatment aimed at killing cancer-cells
Circulating tumour DNA (ctDNA)	is tumor-derived fragmented DNA in the bloodstream that is not associated with cells
Clinical utility	impact on patient outcomes, refers to how likely it is that using the test to guide clinical decisions will significantly improve outcomes related to patients health and well-being (benefits vs. harms, whether using the tests gives added value to not using it, effectiveness, and efficacy). According to EGAPP definitions https://www.cdc.gov/genomics/gtesting/egapp/recommend/method.htm.
Clinical validity	or the strength of clinical correlation, is a test's ability to accurately and reliably identify or *predict* the disorder of interest (sensitivity, specificity, positive predictive value, negative predictive value). According to EGAPP definitions https://www.cdc.gov/genomics/gtesting/egapp/recommend/method.htm
Concordance rate	the number of subjects that are concordant (in agreement) over the total number of subjects assessed

Confidence interval	a type of estimate computed from the statistics of the observed data that proposes a range of plausible values for an unknown parameter (e.g. the mean)
DNA	DNA is an abbreviation for 'deoxyribonucleic acid'. DNA is an acid in the chromosomes in the centre of the cells of living things. DNA determines the particular structure and functions of every cell and is responsible for characteristics being passed on from parents to their children.
End-to-end study	a study that follows patients from testing, through treatment, to final outcomes
Genomics	a term that refers to the molecular composition of a tumor
Immunotherapy	a type of cancer treatment that helps your immune system fight cancer
Indeterminate	results that are neither positive nor negative
Metastatic cancer	a cancer that has spread from its site of origin to other parts of the body
Molecular test	a laboratory test that checks for certain genes, proteins, or other molecules in a sample of tissue, blood, or other body fluid, or that check for certain changes in a gene or chromosome
Multi-gene panel	a genetic test that uses next-generation sequencing to test multiple genes simultaneously
Multiplexing	A method for detecting multiple genetic alteration simultaneously
Mutation (somatic)	a genetic alteration acquired by a cell that can be passed to the progeny of the mutated cell in the course of cell division.
Narrative review	a narrative (non-systematic) which purpose is to identify a few studies that describe a problem of interest
Next Generation Sequencing (NGS)	also called massively parallel, deep sequencing or multigene panel, is a DNA sequencing technology by which entire human genome can be sequenced within a single day
Radiation therapy	a type of cancer treatment that uses high energy beams most often X-rays, but also protons or other types of energy, to kill cancer cells
RNA-based NGS	NGS-method that used ribonucleic acid for the analysis
Gene rearrangement	a programmed DNA recombination event that occurs during cellular differentiation to reconstitute a functional gene from gene segments separated in the genome
Osseous	bony, consisting of bone

PIK3CA gene	a gene that provides instructions for making the p110 alpha (p110α) protein, which is one piece (subunit) of an enzyme called phosphatidylinositol 3-kinase (PI3K)
Polymerase chain reaction	or PCR, is a laboratory method widely used to rapidly make millions to billions of copies (complete copies or partial copies) of a specific DNA sample, in order to study in detail.
Sensitivity	the ability of a test to correctly identify those with the disease (true positive rate)
Single-gene test	a test that looks for changes in only one gene.
Specificity	the ability of the test to correctly identify those without the disease (true negative rate)
Somatic gene mutation	an alteration in DNA that occurs after conception.
Systematic review	a review of the evidence on a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant primary research, and to extract and analyze data from the studies that are included in the review
Turnaround time (TAT)	the total test cycle which includes ordering of test, collection, identification, transportation, preparation, analysis, reporting, interpretation and action
Therapeutic TAT	the interval between when a test is requested to the time a treatment decision is made
Visheral	a term that refers to the internal organs of the body

Appendix 2 Full Search strategy

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions(R) <1946 to October 18, 2021>

Search date: 2021-10-20

1 exp Phosphatidylinositol 3-Kinases/ (42049)

2 (pik3ca or phosphatidylinositol* or phosphoinositide or phospholipids or phosphatidyl or phosphoglycerides or phosphatidyl or phosphoglycerides or phosphoinositides or phosphatidyl or phosphoinositides or ptdIns or phosphatidylisitide).ti,ab,kw,kf. (124714)

3 or/1-2 (144580)

4 exp Breast Neoplasms/ (313669)

5 (ER positive or HER2 negative or ((breast* or mammary or mamma) adj3 (neoplasm* or cancer* or tumor* or tumour* or carcinoma* or onco* or adenocarcinoma* or metastas* or cancerogen*))).ti,ab,kw,kf. (381232)

6 or/4-5 (443035)

7 3 and 6 (5149)

8 limit 7 to yr="2010 -Current" (3806)

Database: Embase <1974 to 2021 October 19>

Search date: 2021-10-20

1 exp Phosphatidylinositol 3-Kinase/ (80914)

2 (pik3ca or phosphatidylinositol* or phosphoinositide or phospholipids or phosphatidyl or phosphoglycerides or phosphatidyl or phosphoglycerides or phosphoinositides or phosphatidyl or phosphoinositides or ptdIns or phosphatidylisitide).ti,ab,hw. (219634)

3 1 or 2 (219634)

4 exp Breast cancer/ (508543)

5 (ER positive or HER2 negative or ((breast* or mammary or mamma) adj3 (neoplasm* or cancer* or tumor* or tumour* or carcinoma* or onco* or adenocarcinoma* or metastas* or cancerogen*))).ti,ab,hw. (665336)

6 4 or 5 (670992)

7 3 and 6 (15111)

8 limit 7 to (embase and yr="2010 -Current") (8675)

Database: Cochrane Database of Systematic Reviews (reviews, protocols), Cochrane Central register of Controlled Trials

Search date: 2021-10-20

#1 MeSH descriptor: [Phosphatidylinositol 3-Kinase] explode all trees 69

#2 (pik3ca or phosphatidylinositol* or phosphoinositide or phospholipids or phosphatidyl or phosphoglycerides or phosphatidyl or phosphoglycerides or

phosphoinositides or phosphatidyl or phosphoinositides or ptdIns or phosphatidylisitide):ti,ab,kw 2856
#3 #1 or #2 2856
#4 MeSH descriptor: [Breast Neoplasms] explode all trees 13847
#5 ((breast* or mammary or mamma) NEAR/5 (neoplasm* or cancer* or tumor* or tumour* or carcinoma* or onco* or adenocarcinoma* or metastas* or cancerogen*)):ti,ab,kw 40090
#6 #4 or #5 40091
#7 #3 and #6 in Cochrane Reviews0
#8 (phosphatidylinositol* or phosphoinositide or phospholipids or phosphatidyl or phosphoglycerides or phosphatidyl or phosphoglycerides or phosphoinositides or phosphatidyl or phosphoinositides or ptdIns or phosphatidylisitide) 2529
#9 #1 or #8 2529
#10 ((ER NEXT positive) or (HER2 NEXT negative) or ((breast* or mammary or mamma) NEAR/3 (neoplasm* or cancer* or tumor* or tumour* or carcinoma* or onco* or adenocarcinoma* or metastas* or cancerogen*))) 40752
#11 #4 or #10 40752
#12 #9 and #11 with Cochrane Library publication date Between Jan 2010 and Dec 2021, in Cochrane Protocols 1
#13 #9 and #11 with Publication Year from 2010 to 2021, in Trials 163
#14 #7 or #12 or #13 with Cochrane Library publication date Between Jan 2010 and Dec 2021 164

Database: Epistemonikos
Search date: 2021-10-20
(title:((pik3ca OR phosphatidylinositol* OR phosphoinositide OR phospholipids OR phosphatidyl OR phosphoglycerides OR phosphatidyl OR phosphoglycerides OR phosphoinositides OR phosphatidyl OR phosphoinositides OR ptdIns OR phosphatidylisitide)) OR abstract:((pik3ca OR phosphatidylinositol* OR phosphoinositide OR phospholipids OR phosphatidyl OR phosphoglycerides OR phosphatidyl OR phosphoglycerides OR phosphoinositides OR phosphatidyl OR phosphoinositides OR ptdIns OR phosphatidylisitide))) AND (title:((breast neoplasm* OR breast cancer* OR mammary OR mamma OR "ER positive" OR "HER2 negative")) OR abstract:((breast neoplasm* OR breast cancer* OR mammary OR mamma OR "ER positive" OR "HER2 negative"))) 2010-2021 (29 systematic reviews)

Database: clinicaltrials.gov
Search date: 2021-10-20
Condition or disease: PIK3CA Mutation: 43

Database: WHO ICTRP
Search date: 2021-10-20
Pik3ca AND breast: 61

Database: Current Controlled Trials
Search date: 2021-10-20 Pik3ca: 7

Appendix 3 List of excluded studies (N=35)

Study First author (reference no.)	Reasons for exclusion
Alvarez 2018 (80)	Not a comparison of test accuracy, but of detection limits. Ineligible population (unclear BC stage).
Andre 2014 (81)	Not a comparison of test accuracy.
Ang 2013(82)	Ineligible population*.
Arnedos 2012 (83)	Not a comparison of test accuracy, and no relevant outcomes
Arsenic 2015(84)	Ineligible population*
Beaver 2014(85)	Ineligible comparison (beween pre- and post surgery samples, not between tests (sanger and ddPCR). Ineligible population. Early stage BC.
Bergholtz 2020(86)	Ineligible population, and ineligible outcome reporting (results for PIK3CA not reported separately).
Bertucci 2016(87)	Ineligible comparison. Not a comparison between diagnostic tests, but between primary tumours and metastases.
Bertucci 2021(88)	Ineligible outcomes. Results for accuracy of tests for detection of PIK3CA not reported separately. Nor are results for our population of interest.
Board 2010(89)	Ineligible comparison. Comparison of ARMs and Scorpion probes but no comparison between tests. Metastatic BC.
Chin 2021(90)	Ineligible comparison. Not a comparison between diagnostic tests.
Dawson 2013(91)	Ineligible comparison. Compares the sensitivity of tests using liquid biopsies (ctDNA) vs tumour cells. MBC.
De Angelis 2020(92)	No full text available
Dirican 2020(93)	Ineligible comparison. No comparison of the PIK3CA detection between tests. Unclear population.
Dumbrava 2021(94)	Not a comparison of the accuracy of tests for detection of PIK3CA mutations. No results by cancer type.
Gerratana 2020(95)	Ineligible population*
Gerratana 2021(96)	Not a comparison of the accuracy of tests for detection of PIK3CA mutations
Harle 2013(97)	Ineligible population*. Neither analytical validity (samples with known mutation status), nor comparison with gold standard.
Higgins 2012(98)	Not a comparison of tests for the detection of PIK3CA mutations.
Hirsch 2018(99)	Ineligible outcomes. Comparisons of platforms, but very few BC patients, and no relevant outcomes for detection of PIK3CA.
Ibrahim 2015(100)	Not a comparison of test accuracy, but about outcome prediction.

Kaur 2019(101)	Ineligible comparison. In silico study.
Lambert 2019(102)	Ineligible population* (breast carcinomas, not mBC). Compares three PCR assays.
Liu 2014(103)	Not a comparison of test accuracy. Review about prognosis.
Markou 2019(44)	Ineligible population*. Unclear if population were ER+/HER2-.
Martinez-Saez 2020(17)	Not a comparison of test accuracy. In silico study.
Ney 2012(104)	Ineligible outcomes. Comparison between two tests, but no results for PIK3CA in patients with mBC.
Ou 2018(105)	Not an original study. Describes various techniques , trial and biomarkers. Review.
Pang 2014(106)	Not a comparison of diagnostic tests, but about prognosis
Raphael 2019(107)	Not a comparison of diagnostic tests. Review about prognosis.
Rothe 2014(108)	Ineligible comparison (plasma vs. biopsies), and not between tests (IonAmpliseq and Illumina technology)
Sobhani 2018(109)	Not a comparison of test accuracy, but about prognosis
Shimoi 2018(110)	Not a comparison of test accuracy, but a comparison between frozen and fresch samples
Song 2013(111)	Ineligible comparison.
Tay 2020(112)	Not an original paper. Not a comparison of test accuracy. Review.
Yang 2019(113)	Not a comparison of test accuracy. Report accuracy of cfDNA and SNV for mutation detection in plasma vs. in tumour tissue.
Zhou 2016(114)	Not a comparison of test accuracy. Compares analyses in tumour tissue and CF.

Appendix 4 Description of the EGAPP quality assessment tool

Table 4.a. Hierarchies of data sources and study designs for the components of evaluation

Level[a]	Analytic validity	Clinical validity	Clinical utility
1	Collaborative study using a large panel of well characterized samples Summary data from well-designed external proficiency testing schemes or interlaboratory comparison programs	Well-designed longitudinal cohort studies Validated clinical decision rule[b]	Meta-analysis of randomized controlled trials (RCT)
2	Other data from proficiency testing schemes Well-designed peer-reviewed studies (e.g., method comparisons, validation studies) Expert panel reviewed FDA summaries	Well-designed case-control studies	A single randomized controlled trial
3	Less well-designed peer-reviewed studies	Lower quality case-control and cross-sectional studies Unvalidated clinical decision rule[b]	Controlled trial without randomization Cohort or case-control study
4	Unpublished and/or non-peer reviewed research, clinical laboratory, or manufacturer data Studies on performance of the same basic methodology, but used to test for a different target	Case series Unpublished and/or non-peer reviewed research, clinical laboratory or manufacturer data Consensus guidelines Expert opinion	Case series Unpublished and/or non-peer reviewed studies Clinical laboratory or manufacturer data Consensus guidelines Expert opinion

[a]Highest level is 1

[b]A clinical decision rule is an algorithm leading to result categorization. It can also be defined as a clinical tool that quantifies the contributions made by different variables (e.g., test result, family history) in order to determine classification/interpretation of a test result (e.g., for diagnosis, prognosis, therapeutic response) in situations requiring complex decision-making.[1]

[1] Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: testing for cytochrome P450 (*CYP450*) polymorphisms in adults with nonpsychotic depression treated with selective serotonin reuptake inhibitors. *Genet Med* 2007;9:819–825.

Table 4.b. Criteria for assessing quality of individual studies (internal validity)[1]

Analytic validity	Clinical validity	Clinical utility
Adequate descriptions of the index test (test under evaluation) Source and inclusion of positive and negative control materials Reproducibility of test results Quality control/assurance measures *Adequate descriptions of the test under evaluation* Specific methods/platforms evaluated Number of positive samples and negative controls tested *Adequate descriptions of the basis for the "right answer"* Comparison to a "gold standard" referent test Consensus (e.g., external proficiency testing) Characterized control materials (e.g., NIST, sequenced) *Avoidance of biases* Blinded testing and interpretation Specimens represent routinely analyzed clinical specimens in all aspects (e.g., collection, transport, processing) Reporting of test failures and uninterpretable or indeterminate results *Analysis of data* Point estimates of analytic sensitivity and specificity with 95% confidence intervals Sample size/power calculations addressed	*Clear description of the disorder/phenotype and outcomes of interest* Status verified for all cases Appropriate verification of controls Verification does not rely on *index test* result Prevalence estimates are provided *Adequate description of study design and test/methodology* *Adequate description of the study population* Inclusion/exclusion criteria Sample size, demographics Study population defined and representative of the clinical population to be tested Allele/genotype frequencies or analyte distributions known in general and subpopulations *Independent blind comparison with appropriate, credible reference standard(s)* Independent of the test Used regardless of test results Description of handling of indeterminate results and outliers Blinded testing and interpretation of results *Analysis of data* Possible biases are identified and potential impact discussed Point estimates of clinical sensitivity and specificity with 95% confidence intervals Estimates of positive and negative predictive values	*Clear description of the outcomes of interest* What was the relative importance of outcomes measured; which were prespecified primary outcomes and which were secondary? *Clear presentation of the study design* Was there clear definition of the specific outcomes or decision options to be studied (clinical and other endpoints)? Was interpretation of outcomes/endpoints blinded? Were negative results verified? *Was data collection prospective or retrospective?* If an experimental study design was used, were subjects randomized? Were intervention and evaluation of outcomes blinded? Did the study include comparison with current practice/empirical treatment (value added)? *Intervention* What interventions were used? What were the criteria for the use of the interventions? *Analysis of data* Is the information provided sufficient to rate the quality of the studies? Are the data relevant to each outcome identified? Is the analysis or modeling explicit and understandable? Are analytic methods prespecified, adequately described, and appropriate for the study design? Were losses to follow-up and resulting potential for bias accounted for? Is there assessment of other sources of bias and confounding? Are there point estimates of impact with 95% CI? Is the analysis adequate for the proposed use?

[1] Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: testing for cytochrome P450 (*CYP450*) polymorphisms in adults with nonpsychotic depression treated with selective serotonin reuptake inhibitors. *Genet Med* 2007;9:819–825.

Table 4.c. Grading the quality of evidence for the individual components of the chain of evidence (key questions)[2]

Adequacy of information to answer key questions	Analytic validity	Clinical validity	Clinical utility
Convincing	*Studies that provide confident estimates of analytic sensitivity and specificity using intended sample types from representative populations* Two or more Level 1 or 2 studies that are generalizable, have a sufficient number and distribution of challenges, and report consistent results One Level 1 or 2 study that is generalizable and has an appropriate number and distribution of challenges	*Well-designed and conducted studies in representative population(s) that measure the strength of association between a genotype or biomarker and a specific and well-defined disease or phenotype* Systematic review/meta-analysis of Level 1 studies with homogeneity Validated Clinical Decision Rule High quality Level 1 cohort study	*Well-designed and conducted studies in representative population(s) that assess specified health outcomes* Systematic review/meta-analysis of randomized controlled trials showing consistency in results At least one large randomized controlled trial (Level 2)
Adequate	Two or more Level 1 or 2 studies that Lack the appropriate number and/or distribution of challenges Are consistent, but not generalizable Modeling showing that lower quality (Level 3, 4) studies may be acceptable for a specific well-defined clinical scenario	Systematic review of lower quality studies Review of Level 1 or 2 studies with heterogeneity Case/control study with good reference standards Unvalidated Clinical Decision Rule (Level 2)	Systematic review with heterogeneity One or more controlled trials without randomization (Level 3) Systematic review of Level 3 cohort studies with consistent results
Inadequate	Combinations of higher quality studies that show important unexplained inconsistencies One or more lower quality studies (Level 3 or 4) Expert opinion	Single case-control study Nonconsecutive cases Lacks consistently applied reference standards Single Level 2 or 3 cohort/case-control study Reference standard defined by the test or not used systematically Studies not blinded Level 4 data	Systematic review of Level 3 quality studies or studies with heterogeneity Single Level 3 cohort or case-control study Level 4 data

[2] Sawaya GF, Guirguis-Blake J, LeFevre M, et al. Update on methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. *Ann Intern Med* 2007;147:871– 875.

Appendix 5 Studies comparing tests for detection of PIK3CA mutations, but with BC populations ineligible for our review (N=6)

Author Year	**Ang 2013**
Title	Novel Method for PIK3CA Mutation Analysis Locked Nucleic AcidPCR Sequencing
Aim	NR
Test comparisons	• PCR and Standard Sanger Sequencing vs. • LNA-PCR Sequencing Assay vs. • PCR-Mass Spectroscopy
Population/no of samples tested	60 FFPE breast tissue samples
Type of BC	31 preneoplastic breast lesions, 21 breast carcinomas (including in situ, invasive, and *metastatic*), and 8 normal breast samples
Outcomes	Consistency/discordance between test results
Study design	Retrospective
Selection of patients	NR
Country	USA
Results	Four of 19 mutated cases were not picked up with PCR-MS, 15 of 19 mutated cases were not picked up with Sanger sequencing, and all 19 mutated cases were detected with LNA-PCR.
Author Year	**Arsenic 2015**
Title	Comparison of targeted next-generation sequencing and Sanger sequencing for the detection of PIK3CA mutations in breast cancer
Aim	To evaluate the concordance between NGS and SGS for the most important hotspot regions in exon 9 and 20, to investigate additional hotspots outside of these exons using NGS, and to correlate the PIK3CA mutation status with the clinicopathological characteristics of the cohort.
Test comparisons	• NGS (PIK3CAsemiconductor next-generation sequencing) using PCR vs. • Sanger sequencing
No of patients/ samples tested	N=186
Type of BC	Primary breast cancer (BC); no information on molecular sub-types
Outcomes	Concordance
Study design	Retrospective
Patient selection	NR
Country	Germany
Results	64 tumors had PIK3CA mutations, (55 of these mutations occurred in exons 9 and 20). Out of these 55 of 66 tumours had PIK3CA mutations with NGS, of which 52 that also were detected by Sanger sequencing resulting in a concordance of 98.4 % between the two sequencing methods. The three mutations missed by SGS had low variant frequencies below 10 %. Additionally,4.8 % of the tumors had mutations in exons 1, 4, 7, and 13 of PIK3CA that were not detected by SGS.
Author Year	**Gerratana 2020**
Title	Performance of a novel Next Generation Sequencing circulating tumor DNA (ctDNA) platform for the evaluation of samples from patients with metastatic breast cancer (MBC)
Aim	To explore the performance of the novel diagnostic NGS platform PredicinePLUS™ and to compare its results with the clinically available Guardant360™ platform for possible analytical inconsistencies.
Test comparisons	• PredicinePLUS™ vs. • Guardant360™ platform

No of patients/ samples tested	N=15
Type of BC	MBC; unknown molecular sub-types
Outcomes	Concordance
Study design	Retrospective
Patient selection	Patients treated at the Robert H. Lurie Comprehensive Cancer Center at Northwestern University of Chicago (IL, USA) between September 2016 and August 2017, under the single Institution Investigator Initiated Prospective Trial NU16B06
Country	USA
Results	Cohen's kappa: 1.0 (no measure of dispersion): 8 WT, and 6 mutated samples (42.8% mutated)
Author Year	**Markou 2019**
Aim	Nuclease-Assisted Minor Allele Enrichment Using Overlapping Probes-Assisted Amplification-Refractory Mutation System: An Approach for the Improvement of Amplification-Refractory Mutation System-Polymerase Chain Reaction Specificity in Liquid Biopsies
Test comparisons	• NAPA: NaME-PrO-assisted ARMS-PCR vs. • Allele-specific, asymmetric rapid PCR, and melting analysis (described under Multiplex-PCR and Melting Analysis) vs. • Droplet Digital PCR (ddPCR)-
No of patients/ samples tested	41 samples from breast cancer patients with clinically confirmed metastasis; 20 control samples
Type of BC	MBC; unknown molecular subtypes
Outcomes	Concordance
Study design	Retrospective
Patient selection	NR
Country	Authors from the US and Greece. Unclear what country the were retrieved from.
Results	(i) Novel NAPA assay vs ddPCR The concordance between tests for detection of PIK3CA E545 K hotspot mutation in EpCAM-positive CTC fraction was 22/24 (91.2%) (p=0.022, chi-square test) The concordance between tests for detection of PIK3CA E545 K hotspot mutation in cfDNA was 15/17 (88.2%) (p=0.003, chi-square test) In total 37/41 (90.2%) were concordant across the two tests Discordance: 4 samples: 2 pos with NAPA but negative with ddPCR; 2 pos with ddPCR but neg with NAPA (ii) Allele specific, assymetric rapid PCR and melting analysis assay vs ddPCR The concordance between tests for detection of PIK3CA E545 K mutation in EpCAM-positive CTC fraction was 22/24 (92.7%) (19 samples were found negative, and 3 positive by both methods. Discordance: 2 samples: One sample was positive by our previous ultrasensitive assay and negative by ddPCR, and one sample was negative by the ultrasensitive assay and positive by ddPCR. The concordance between tests for detection of PIK3CA E545 K mutation in cfDNA was 12/17 (70.6%) for samples from confirmed MBC patients (9 samples were found negative by both assays, and three samples were positive by both assays). Discordance: 5 samples: In total 34/41 samples were concordant (82.9%) across tests
Author Year	**Harle 2013**
Title	Analysis of PIK3CA exon 9 and 20 mutations in breast cancers using PCR-HRM and PCR-ARMS: Correlation with clinicopathological criteria
Aim	To evaluate the relationship between PIK3CA exon 9 and 20 mutations and conventional clinicopathological criteria and to compare

	the sensitivity of the two techniques by examining the correlation of the results achieved using both methods.
Test comparisons	• PCR amplification-refractory mutation system Scorpions® (ARMS) vs. • High-Resolution Melting PCR assay (HRM)
No of patients/ samples tested	102 breast tumor samples
Type of BC	Invasive BC (mixed types, stage, and hormonal status/molecular subtypes)
Outcomes	Consistency (and detection sensitivity)
Study design	Retrospective
Selection of patients	149 invasive breast carcinoma tumor specimens, from patients diagnosed between 2008 and 2009, were included retrospectively with informed consent. Specimens were collected as AFA-fixed PE tissues from the institutional Biobank.
Country	France
Results	PCR-HRM: PIK3CA mutations detected in 28 tumours (27.5%) PCR-ARMS: PIK3CA mutations detected in 23 tumours (22.5%) Discordant samples: 5 PCR-ARMS was more sensitive than PCR-HRM (sensitivity 0.5 and 5-10% of mutated DNA, respectively).
Author Year	**Lambert 2019**
Title	Comparison of Three Real-Time PCR Assays for the Detection of PIK3CA Somatic Mutations in FFPE Tissues of Patients with Breast Carcinomas
Aim	To compare three assays for the somatic mutation detection of PIK3CA gene in FFPE tissues of patients with breast cancer
Test comparisons	• Cobas® PIK3CA Mutation Test (Roche Diagnostics, Meylan, France), • PCR amplification-refractory mutation system Scorpions® (ARMS) and High-Resolution Melting PCR assay (HRM)
Confirmation test	Discrepant samples were confirmed using NGS.
No of patients/ samples tested	46 FFPE breast carcinomas samples of patients under treatment
Type of BC	BC (type, stage, and hormone status/molecular subtypes not described)
Patient demographics	No information
Outcomes	Concordance between tests (Cohen's kappa)
Study design	Retrospective
Patient selection	Breast carcinoma samples from patients treated for BC at the Institut de Cancérologie de Lorraine were retrospectively collected, with the informed consent of patients.
Country	France
Results	No of PIK3CA mutations among 46 samples across tests: Cobas®: 17 (37.8%), ARMS: 13 (28.36%) and HRM assays: 19 (41.3%) Cohen's kappa: Cobas® and HRM: 0.95[0.86;1], Cobas® and ARMS: 0.75[0.55;0.95] HRM and ARMS: 0.72[0.51;0.92] Five samples had discrepant results. Three different PIK3CA mutations were detected in one sample.

Appendix 6 Quality of evidence- results of the EGAPP tool

	Nteliopaulus 2021	O'Leary 2019	Zivanovich 2020
Level of study Hierarchy – (grade 1-4)	3	1	1
ANALYTICAL VALIDITY			
Adequate description of index test			
Source and inclusion of positive and negative control materials	NR	NR	Yes, for ddPCR
Reproducibility of test results	NR	yes	NR
Quality control/assurance program	NR	NR	NR
Adequate description of the test under evaluation			
Specific methods/platforms evaluated	Yes	Yes	yes
Number of positive samples and negative controls tested	NA	NA	NA
Adequate description of the terms for the right answer			
Comparison to a "gold standard" referent test	no	no	no
Consensus (e.g., external proficiency testing)	NR	NR	NR
Characterized control materials (e.g., NIST, sequenced)	NA	NA	Yes, for ddPCR
Avoidance of biases			
Blinded testing and interpretation	NR	yes (testing performed at separate laboratories)	NR
Specimens represent routinely analyzed clinical specimens in all aspects (e.g., collection, transport, processing)	Yes (add here)	No (all were trial participants)	NR
Reporting of test failures and uninterpretable or indeterminate results	Yes (7 indeteriminate results, and 5 samples excluded)	No	Yes (information in supplementary information)
Analysis of data			
Point estimates of analytic sensitivity and specificity with 95% CI	No	No	No
Sample size/power calculations addressed	No	No	No
CLINICAL VALIDITY			
Clear description of the disorder/phenotype and outcomes of interest			
Status verified for all cases	No (30 of 96 samples tested with both NGS platforms)	No (27 samples were excluded)	No (162 of a total 363)
Appropriate verification of controls	NA	NA	Yes, for ddPCR
Verification does not rely on index test result	Unclear	Unclear	Unclear
Prevalence estimates are provided	No	No	No
Adequate description of study design and test/methodology			

	Partly, more information that in some of the other studies	No	Yes
Adequate description of the study population			
Inclusion/exclusion criteria	No (apart from cancer type and stage)	Partly (no exclusion criteria)	Partly (no exclusion criteria)
Sample size, demographics	No	No	Yes
Study population defined and representative of the clinical population to be tested	No	yes (unselected patients)	Yes (consecutive clinical patients)
Allele/genotype frequencies or analyte distributions known in general and subpopulations	NA	NA	NA
Independent blind comparison with appropriate, credible reference standard(s)			
Independent of the test	unclear	unclear	unclear
Used regardless of test results	Yes (only 35 of 96 samples were tested with both methods due to limited cfDNA; 5 were excluded)	Yes (all samples were tested)	Unclear (162 of 363 samples were tested)
Description of handling of indeterminate results and outliers	Unclear (Mention inderterminate results, but unclear how these were managed)	no	No
Blinded testing and interpretation of results	Unclear, CLIA	Unclear	Unclear
Analysis of data			
Possible biases are identified and potential impact discussed	Partly	Yes	Partly
Point estimates of clinical sensitivity and specificity with 95% CI	No, report concordance with 95% CI	No, report concordance with 95% CI	No, report concordance, but no CI
Estimates of positive and negative predictive value	no	no	no
Clinical utility- No data available			
Analytical validity			
-Inadequate	X	X	X
Clinical validity			
-Convincing			
-Adequate			
-Inadequate	X	X	X

CI: Confidence Interval; RS: Reference standard; NA: data/information not available; ROC: Receiver Operation Characteristics curve:

Appendix 7 Quality of reporting: results of the STARD checklist

STARD items	Nteliopoulos 2021	O'Leary 2019	Zivanovich 2020
1	Yes, Indicated in the abstract	Yes, Indicated in the abstract	Yes, Indicated in the abstract
2	Yes, but do not mention study design	Yes, but do not mention study design	Yes
3	Yes	Partly	Yes
4	Partly, (objectives, but no hypothesis)	Partly, (objectives, but no hypothesis)	Partly (objectives but no hypothesis)
5	NR	NR	NR
6	NR	NR	Yes
7	NR	Included in trial	Included in trial
8	yes	NR	yes
9	NR	unselected	consecutive
10a	unclear	unclear	yes
10b	unclear	unclear	yes
11	no	no	no
12a	NR	NR	yes
12b	NR	NR	yes
13a	NR	blinded	NR
13b	NR	blinded	NR
14	yes	yes	Yes, Spearman's correlation
15	Unclear (7 indeterminate and 5 excluded)	NR	NR
16	NR	NR	NR
17	unclear	unclear	NR
18	NR	NR	NR
19	yes	NR	yes
20	no	NR	yes
21a	NR	NR	NA
21b	NR	NR	NA
22	no	NR	Testing done in parallell
23	yes	yes	Yes, information in Supplement
24	no	yes	Yes, but no CI , and not for PIK3CA separately
25	NR	NR	NR
26	yes	yes	Partly
27	yes	yes	Partly
28	yes	No	Yes
29	no	No	No
30	Report (partly commercial) funding, and possible financial CoI. Role of funders **not** described.	Report funding, but no financial CoI. Role of funders idescribed.	Report (partly commercial) funding, and possible financial CoI. Role of funders described

STARD checklist items:
TITLE or ABSTRACT
1 Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)
ABSRACT
2 Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)
INTRODUCTION
3 Scientific and clinical background, including the intended use and clinical role of the index test
4 Study objectives and hypotheses
METHODS
Study design
5 Whether data collection was planned before the index test and reference standard were performed (prospective) or after (retrospective study)

Participants
6 Eligibility criteria
7 On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)
8 Where and when potentially eligible participants were identified (setting, location and dates)
9 Whether participants formed a consecutive, random or convenience series
Test methods
10a Index test, in sufficient detail to allow replication. Note: All three phases should have been described: pre-analytical, analytical, post-analytical.
10b Reference standard, in sufficient detail to allow replication
11 Rationale for choosing the reference standard (if alternatives exist)
12a Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory
12b Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory
13a Whether clinical information and reference standard results were available to the performers/readers of the index test
13b Whether clinical information and index test results were available to the assessors of the reference standard
Analysis
14 Methods for estimating or comparing measures of diagnostic accuracy
15 How indeterminate index test or reference standard results were handled
16 How missing data on the index test and reference standard were handled
17 Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory
18 Intended sample size and how it was determined
RESULTS
Participants
19 Flow of participants, using a diagram
20 Baseline demographic and clinical characteristics of participants
21a Distribution of severity of disease in those with the target condition
21b Distribution of alternative diagnoses in those without the target condition
22 Time interval and any clinical interventions between index test and reference standard
Test results
23 Cross tabulation of the index test results (or their distribution) by the results of the reference standard
24 Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)
25 Any adverse events from performing the index test or the reference standard
DISCUSSION
26 Study limitations, including sources of potential bias, statistical uncertainty, and generalisability
27 Implications for practice, including the intended use and clinical role of the index test
OTHER INFORMATION
28 Registration number and name of registry
29 Where the full study protocol can be accessed
30 Sources of funding and other support; role of funders

Appendix 8 Funding, role of funders and conflicts of interest

Author Year	Financial support	Role of funders	Conflict of interest	Comments
Original studies (N=3)				
Nteliopaulus 2021(39)	Supported by sequential programme grant funding from Cancer Research UK to J.A. Shaw and R.C. Coombes, a Clinical and Translational Research Committee Award (C14315/A13462) and a Science Committee Programme Award (C14315/A23464) and by Inivata Ltd. G.	NR	KH, WE, EG and NR are current or former employees, officers, consultants and/or share-holders of Inivata Ltd or Inivata Inc. R.C. Coombes reports speaker frees from Pfzer and has shares in Carrick Ltd. L. Kenny reports receiving speakers bureau honoraria from Pfzer, and is a consultant/advisory board member for Celgene, Novartis, and Lilly. J. Stebbing is the Editor-in-Chief of Oncogene, sat on SABs for Celltrion, Vor Biopharma and Benevolent AI, and Chairs the Board of Directors for BB Biotech Healthcare Trust and Xerion Healthcare. All the other authors declare that they have no competing interests. Nteliopoulos is funded by the support from the Imperial NIHR Biomedical Research Centre Programme	Partly commercially funded.
O'Leary 2019 (40)	The Medical Research Council (MR/N002121/ 1), The Breast Cancer Now Toby Robins Research Centre with support from the Mary-Jean Mitchell Green Foundation and Pfizer. National Institute for Health Research provided funding to the Royal Marsden and Institute of Cancer Research Biomedical Research Centre. B. O'Leary, funding from Pfizer to institution; F. Andre, funding from AstraZeneca, Novartis, Pfizer, Lilly, Roche to institution; S. Loibl, funding from Pfizer, Roche, Celgene, Amgene, Novartis, Abbvie, AstraZeneca, Seattle Genetics, Teva, Vifor Pharma to institution; S. Loi, funding from Roche/Genentech, Pfizer, Novartis, Merck, Puma Biotechnology, Bristol-Myers Squib to institution; N. Turner, AstraZeneca, funding from Pfizer, Roche, and BioRad to institution.	The funding organizations played no role in the review and interpretation of data, preparation of manuscript, or final approval of manuscript. The funding organizations played a direct role in the design of study and choice of enrolled patients.	**Stock Ownership**: J. Jiang, Pfizer; Y. Liu, Pfizer; XH, Pfizer; C. Huang Bartlett, Pfizer. Honoraria: S. Loibl, Pfizer, Roche, AbbVie, Amgen, AstraZeneca, Celgene, Novartis, Seattle Genetics, Teva, Vifor, Prime, Daiichi; M. Cristofanilli, Dompe´ Farmaceutici, Pfizer **Other Remuneration**: F. Andre, travel, accommodation, and expenses, Novartis, Roche, GlaxoSmithKline, AstraZeneca **Employment or Leadership**: J. Jiang, Pfizer; Y. Liu, Pfizer; C. Huang Bartlett, Pfizer; S. Loibl, GBG, BIG, ESMO Breast. **Consultant or Advisory Role**: S. Loibl, Pfizer, Roche, Novartis, Seattle Genetics, AbbVie, Amgen, AstraZeneca, Celgene; S. Loi,AstraZeneca/MedImmune, Seattle Genetics, Bristol-Myers Squibb, Pfizer, Novartis, Roche/Genentech, Merck Sharp & Dohme; M. Cristofanilli, Dompe´	Partly commercially funded.

			Farmaceutici, Newomics, Vortex Biosciences; NC, Roche, Pfizer, Novartis, AstraZeneca, CytoDyn, Merus.	
Zivanovich 2020 (41)	S-J.D and A.Z.B. received funding for this research. S-J.D received funding through the National Health and Medical Research Council of Australia (grant number APP1085014, https://www.nhmrc.gov.au), the Peter MacCallum Cancer Centre Women's Cancer Research Program (https://www.petermac.org), **Genentech (https://www.gene.com),** and the Australian Cancer Research Foundation (https://www.acrf.com.au). S-J.D. was supported by a National Breast Cancer Foundation (https://nbcf.org.au) and Victorian Cancer Agency (http://victoriancanceragency.vic. gov.au) Fellowship and a CSL Centenary Fellowship (https://www.cslfellowships.com.au). A.Z.B. was supported by an Australian Postgraduate Award administered by the University of Melbourne (https://www.unimelb.edu.au) and a PhD Top Up Scholarship administered by **Cancer Therapeutics CRC**, Melbourne, Australia (https://cancercrc.com).	The funders played no role in the study design, data collection and analysis, decision to publish, or the preparation of the manuscript.	M.A.D. has been a member of advisory boards for CTX CRC, Storm Therapeutics, Celgene, Cambridge Epigenetix, and GSK. S. L. receives research funding to her institution from Novartis, Bristol Meyers Squibb, Merk, Roche-Genentech, Puma Biotechnology, Pfizer, and Eli Lilly. She has acted as a consultant (not compensated) to Seattle Genetics, Pfizer, Novartis, BMS, Merck, AstraZeneca, and RocheGenentech. She has acted as a consultant (paid to her institution) to Aduro Biotech. S.Q.W. has received travel support from BioRad. S-J. D. has received research funding to her institution from Roche-Genentech and CTX CRC. She has been a member of advisory boards for AstraZeneca.	Partially commercially funded.
Narrative reviews (N=1)				
Fusco 2021(42)	The authors declare that this study received funding from Novartis Farma SpA.	The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.	NR	Commercial funding.

Appendix 9 Estimated cost related to PCR and NGS testing, St. Olav's University Hospital and Oslo University Hospital

RT-PCR

Estimated cost for testing with RT-PCR, St Olav's University Hospital

	St. Olav's University Hospital	
	1 patient (NOK)	**10 patients (NOK)**
Reagent cost	1,200	1,200
Personal cost	930 (bioengineer: 810, pathologist: 120)	364 (bioengineer: 244, pathologist: 120)
Sum	**3,113**	**1,419**

NGS

Estimated cost for testing with NGS, St Olav's University Hospital and Oslo University Hospital

	St. Olav's University Hospital*		Oslo University Hospital** (NOK)
	1 patient (NOK)	**10 patients (NOK)**	
Reagent cost	17,580	2,580	7,168
Personal cost	1,462 (bioengineer: 948.70, molecular biologist: 333.33 pathologist: 179.49)	705 (bioengineer: 192.30, molecular biologist: 333.33 pathologist: 179.49)	3,800 (bioengineering:3,000, Pathologist: 800)
Sum	**19,042**	**3,285**	**10,968**

**Ion Torrent S5 (personal communication by Liv Solvår Nymark, St. Olav's University Hospital)*

***Ion Torrent S5, Oncomine Childhood Cancer Research Assay (personal communication by Martin Andreas Furu, Oslo University Hospital)*

Appendix 10 Activity log

Activity	Date
Suggestion submitted (ID2019_070)	07.05.2019
HTA commissioned	26.08.2019, and updated 14.12.2020 (ID2019_070)
External experts pointed out by RHA	20.08.2020
Work paused due to prioritisation of ROS1 report	01.09.2020
Updated search for literature	22.10.2021
Report sent for expert review	23.02.2022
Internal approval at NIPH	18.03.2022
Report submitted	18.03.2022

NIPH

Published by the Norwegian Institute of Public Health
Mars 2022
P.O.B 4404 Nydalen
NO-0403 Oslo
Phone: + 47-21 07 70 00
The report can be downloaded as pdf
at www.fhi.no/en/publ/