Cancer screening in Europe

Rapid review 3

What is the evidence from recent trials and reviews for the efficacy, harm-benefit and cost-effectiveness of new technologies in cancer screening and early diagnosis?

To download the latest version of this document, together with the full Evidence Review Report that it informed, please visit:

www.sapea.info/cancerscreening

SAPEA
Science Advice for Policy by European Academies

- DOI: ~forthcoming
- Downloadable from https://www.sapea.info/cancer-screening/

Version history

Version	Date	Summary of changes
1.0	2 March 2022	First published version

Publisher

SAPEA
c/o acatech
Pariser Platz 4a
10117 Berlin, Germany

Contact

SAPEA Communications Office
Rue d'Egmont 13
1000 Brussels, Belgium
contact@sapea.info

SAPEA, Science Advice for Policy by European Academies. (2022). *Cancer screening in Europe: Rapid review 3*. Berlin: SAPEA.

SAPEA
Science Advice for Policy by European Academies

Rapid Review 3

What is the evidence from recent trials and reviews for the efficacy, harm-benefit and cost-effectiveness of new technologies in cancer screening and early diagnosis?

Rapid Review Details

Review conducted by:
A team led by the Specialist Unit for Review Evidence (SURE) for SAPEA (Science Advice for Policy by European Academies)

Review Team:
- Dr Hui-Ling Ou, Research Associate, Cambridge Centre Lung Cancer Early Detection Group, University of Cambridge, UK
- Dr Nicholas Courtier, Senior Lecturer Radiography, School of Healthcare Sciences, Cardiff University, UK
- Dr Alison Weightman, Director Specialist Unit for Review Evidence, Cardiff University, UK
- Louise Edwards, Hub Manager, Academia Europaea, Cardiff University, UK

Method:
This is one of three rapid reviews - a lighter form of a full systematic review that takes account of time constraints. The top-line results are included in the main SAPEA Evidence Review Report, with cross-referencing between the documents.

The review summarises a valuable subset of the evidence base, emphasising the findings from recent randomised and other controlled clinical trials. This review includes diagnostic accuracy studies carried out within controlled trials published since 2017, supplemented with data from published systematic reviews of case-control and diagnostic accuracy studies. To meet deadlines, a pragmatic and precise search strategy was employed; it is possible that further controlled trials would have been identified if there had been time for a detailed and sensitive systematic search. The timeline also precluded any statistical or meta-analysis of findings unless these were available from published systematic reviews. No formal critical appraisal was carried out although information is provided on whether the trial included a power calculation. Data extraction and summary were undertaken by different reviewers and, although reviewed by another author, these have not been independently checked for accuracy and consistency.

Acknowledgements:
The advisory group for the review team, comprising the Chairs of the Expert Workshop, Professor Ole Petersen (Academia Europaea), members of SAPEA, the Group of Chief Scientific Advisors (Advisors) and the SAM Unit. Professor Adrian Edwards (Professor of General Practice, Cardiff University) and Professor Jacek Jassem (Head, Department of Oncology & Radiotherapy, Medical University of Gdańsk) for reviewing and commenting on the draft pre-publication.

Disclaimer: The authors of this work declare that they have no conflicts of interest.

'What is the evidence from recent trials and reviews for the efficacy, harm-benefit and cost-effectiveness of new technologies in cancer screening and early diagnosis??

TOPLINE SUMMARY

Who is this summary for?

To support the work of SAPEA in providing evidence to the European Commission's Group of Chief Scientific Advisors on cancer screening in Europe.

Background

This review is one of three rapid reviews conducted on the topic of cancer screening in Europe. It was produced specifically for the expert workshop convened to discuss the main scientific elements to consider, and best practices to promote, for optimising risk-based cancer screening and early diagnosis throughout the EU. This final version has been revised to address feedback received on earlier drafts and supplements the workshop report (available on the SAPEA website).

Aim

To examine the published evidence base for the question: 'What is the evidence from recent trials for the efficacy, harm-benefit and cost-effectiveness of new technologies in cancer screening and early diagnosis?'.

Rapid review method

A literature search was conducted in October 2021 for diagnostic accuracy studies carried out within controlled trials published since 2017, supplemented with data from published systematic reviews of case-control and diagnostic accuracy studies. Trials and systematic reviews were included if they examined new technologies (including artificial intelligence [AI], imaging and biomarkers) in screening for first diagnosis of any cancer, and included data on efficacy, harm-benefit or cost-effectiveness.

Key findings

Biomarkers [11 studies within trials and 11 systematic reviews of diagnostic studies]:

- Biomarker panels show better specificity in cancer detection than single markers.
- Biomarkers not only facilitate cancer detection, but also enhance detection of pre-cancerous lesions, e.g., Cytosponge®-TFF3 and saliva cytokines.
- Across various cancer types, the biomarkers for colorectal cancer screening are the most intensively studied, including genomic, epigenetic and protein markers detected in blood, stool, urine and tissue.

Imaging and artificial intelligence [8 studies within trials and 1 systematic review of diagnostic studies]:

- Novel image-enhanced endoscopy can improve early detection of upper GI-tract lesions in high-risk populations.
- There is small-scale evidence for superiority of blue light imaging in bright mode over linked colour imaging in detection of colorectal adenomas but this requires confirmation.

- Retrospective evidence, and lack of prospective evidence, suggest that current AI is not sufficiently specific to replace double radiologist reading in breast screening programmes.

Strength of evidence

The evidence is derived from studies embedded within controlled clinical trials and systematic reviews of case-control or diagnostic test accuracy studies.

Full report

Table of Contents

1. Background

This Rapid Review is one of three reviews being conducted to support the work of Expert Groups convened to assist the European Commission Scientific Advice Mechanism (SAM) in developing policy guidance in relation to cancer screening. As described in the Scoping Paper[1], this review supports the third of the Expert Group workshops convened to discuss the question **"Which are the main scientific elements to consider, and best practices to promote, for optimising risk-based cancer screening and early diagnosis throughout the EU?"**

An advisory group was formed to provide guidance to the review team, comprising the Chairs, Professor Ole Petersen (Academia Europaea), members of SAPEA, the Group of Chief Scientific Advisors and the SAM Unit.

1.1 Purpose of this review

Following detailed discussions with the advisory group, the specific question for the rapid review to inform the third workshop was:

"What is the evidence from recent controlled trials for the efficacy, harm-benefit and cost-effectiveness of new technologies in cancer screening and early diagnosis?"

Following completion of the search, it was subsequently agreed to include published systematic reviews of case-control and test-accuracy studies, given the large amount of evidence summarised within these reviews.

1.2 Research question

Rapid Review Question
What is the evidence from recent controlled trials, and systematic reviews of case-control and test accuracy studies, for the efficacy, harm-benefit and cost-effectiveness of new technologies in cancer screening and early diagnosis?

2. Results

2.1 Summary of the evidence base

In all, 19 studies included within trials and 12 systematic reviews of case control/diagnostic accuracy studies have been summarised. We provide a narrative overview of the identified evidence below under two headings:

- Biomarkers

[1] Scientific Advice Mechanism. European Commission's Group of Chief Scientific Advisors. Scoping Paper: Cancer Screening. 22 April 2021

- Imaging & Artificial Intelligence

A summary of each included study is provided in Section 2.2.

Biomarkers

Data from 11 trial reports and 11 systematic reviews (of case control/diagnostic accuracy studies) were included. Information from studies within individual trials or systematic reviews of multiple case-control or diagnostic accuracy studies were extracted and summarised in the Table (Section 2.2). Most studies focused on identifying biomarkers for cancer early diagnosis using liquid biopsies, which can be further divided into protein biomarkers, epigenetic biomarkers, DNA (circulating tumour DNA and mitochondrial DNA) and extracellular RNAs.

Specific protein or antibody biomarkers have been reported useful for discriminating between cancer patients and cancer-free controls. Some markers are specific for cancer type, e.g., serum IDH1 level for non-small cell lung cancer (Sun et al., 2020) and pepsinogen for digestive tract cancers (In et al., 2021; Kunzmann et al., 2018), whilst some markers are mutually shared among different cancers, e.g., p53 antibody for lung and colorectal cancer (Harlid et al., 2021; Sullivan et al., 2021). In addition to cancer detection, a particular biomarker - TFF3 used together with a special specimen collection device - Cytosponge® has shown promising effect on early diagnosis of Barrett's oesophagus, which is the pre-cancerous lesion of oesophageal adenocarcinoma (Fitzgerald, di Pietro, O'Donovan, Maroni, et al., 2020; Fitzgerald, Di Pietro, O'Donovan, Muldrew, et al., 2020; Swart et al., 2021). To mitigate invasive procedures during screening, levels of several cytokines in saliva could be of good use for risk stratification in oral cancer screening (Chiamulera et al., 2021).

Plasma DNA and the methylation frequency are also widely studied across different cancer types, including breast (Sturgeon et al., 2017; Sturgeon et al., 2021), colorectal (Anghel et al., 2021), lung (Hubers et al., 2017; Liu et al., 2017) and melanoma (Guo et al., 2019). Specifically, the methylation status of the adenomatous polyposis coli (APC) promoter has been shown correlated with higher risk of gastric cancer, based on a meta-analysis pooling data from 8 RCTs. Higher incidence of APC was observed in tissues and blood of patients with gastric cancer (OR 3.86; 95%CI 1.71-8.74; P = 0.001) compared with patients without (Zhou et al., 2020). In general, a biomarker panel demonstrated higher specificity for cancer detection than single marker. Similar to protein biomarkers, methylation levels of specific genes were shared among different cancer types, which may be used for general cancer screening.

Several studies also evaluated the efficiency of extracellular RNAs such as miRNA, lncRNA or circular RNA in screening of colorectal, gastric, oesophageal, lung and ovarian cancer (Chu et al., 2018; Hulstaert et al., 2021; Saheb Sharif-Askari et al., 2020; Yu et al., 2020). One miRNA family, miR-21, was found associated with worse overall survival of colorectal cancer (Saheb Sharif-Askari et al., 2020). Panels of miRNA rendered better sensitivity and specificity in detecting lung and ovarian cancer (Chu et al., 2018; Hulstaert et al., 2021).

One study reported the efficiency of DNA quantitative cytology in detecting endometrial cancer (Yang et al., 2019). In terms of kidney cell carcinoma, a recent meta-analysis across 6 RCTs revealed that liquid biomarkers, e.g., miRNAs, proteins and metabolites in urine or plasma, might not be

ready for clinical integration of cancer diagnosis, for which more validation is required (Campi et al., 2021).

As far as the invasiveness is concerned, urinary biopsy is among the least invasive diagnostic tools. One recent systematic review pooled results across 13 RCTs and assessed the quality of using urinary volatile organic compounds (VOCs) for cancer early detection. Despite distinctive VOC profiles in different cancer types (prostate cancer, gastrointestinal cancer, leukaemia/lymphoma, lung cancer and bladder cancer) and promising performance, inconsistencies across RCTs undermine the application of such method to broader populations (Wen et al., 2020).

In summary, there is consistency across various studies embedded in trials, in relation to biomarkers, but the small size of validation groups and heterogenicity of population included per trial may limit the extrapolation of results. Biomarker panels tend to show better specificity in cancer detection than single markers. (Anghel et al. 2021; Carozzi et al. 2017 a/b; Chu et al. 2018; Hulstaert et al. 2021; Tarney et al. 2019).

Two studies reported cost-effectiveness of biomarkers in cancer screening in very different settings. Sullivan et al (2020) found that the cost per stage I/II lung cancer detected using the autoantibody test within EarlyCDT after 2 years was £116,000. In relation to Cytosponge use for Barrett's oesophagus, an additional 0.015 QALYs per patients was generated with the Cytosponge®-TFF3 screening, rendering an incremental cost-effectiveness ratio (ICER) of £5500 per QALY gained (Fitzgerald, di Pietro, O'Donovan, Maroni, et al., 2020; Fitzgerald, Di Pietro, O'Donovan, Muldrew, et al., 2020; Swart et al., 2021).

Imaging and artificial intelligence

Seven studies of novel imaging technologies were identified. Five RCTs focused on early detection of gastrointestinal cancers (Dohi et al. 2019; Gao et al., 2021; Gruner et al. 2021; Ferreira et al. 2021; Yoshida et al. 2021) , one on improved discrimination/management of skin malignancy (Ferrándiz et al. 2017), and a modelling study that sought to personalise lung cancer risk in a large-scale RCT-derived cohort (Hostetter et al. 2017). All study populations were of average to high risk for the target cancer. Two populations were European (Ferrándiz et al. 2017; Gruner et al. 2021). A further, single arm trial, is also summarised in the text (Chauvie et al. 2020).

One systematic review of AI image analysis in breast cancer screening was included, drawing together the evidence from 12 test accuracy studies (Freeman et al. 2021).

GI tract cancers: Overall, improved **detection efficacy** for early gastric and oesophageal lesions was demonstrated for image-enhanced endoscopy using magnification plus narrow-band imaging (NBI) or laser light techniques – light linked colour imaging (LCI) or blue laser imaging (BLI) in bright mode – compared to standard white light imaging (WLI). Evidence for the upper GI tract sites is moderate to high based on consistency of RCT results. **Mortality** and **cost-effectiveness** outcomes were not reported.

Gastric cancer: The largest trial (Yoshida et al. 2021) reported comparable detection rates for second generation NBI versus WLI but with slightly improved positive predictive value for NBI, suggesting potential to reduce false-positive results. Smaller RCTs found BLI-bright (Dohi et al. 2019) and LCI (Gao et al. 2021) had significantly improved detection rate compared to WLI (across disease stage and histology).
Oesophageal cancer: A French RCT (Gruner et al. 2021) found NBI was more specific (80% vs 66%) and sensitive (38% vs 21%) than Lugol chromoendoscopy for the detection of squamous cancers. Supplemental NBI imaging could improve the detection of early neoplasia.
Colorectal cancer: A small Brazilian study (Ferreira et al. 2020) reported a significantly higher adenoma detection rate for LCI (68%) versus both WLI (56%) and BLI-bright (56%); including a superior flat-lesion detection rate.
Skin malignancy: The addition of tele-dermoscopic images to standard clinical images improved the accuracy index (correct decisions percentage) for suspicious skin lesions from 79.2% to 94.3% (Ferrándiz et al. 2017). This higher accuracy made tele-dermoscopy the dominant strategy, with a significantly lower cost-effectiveness ratio.
Lung cancer: Incorporating three clinically-accessible personalised data points – smoking history, sex, nodule location – to an established non-personalised risk model enabled improved malignancy risk predictions and follow-up recommendations to be made (Hostetter et al. 2017).
Another trial, Chauvie et al (2020), reported that AI may facilitate conventional screening but this was a single arm trial (the SOS study) so it does not strictly meet the criteria for inclusion in this review.

The **systematic review** (Freeman et al. 2021) tested accuracy of standalone AI algorithms or AI-assisted radiologists to detect breast cancer in digital mammogram screening or test sets. Cancer type (e.g. grade, stage, prognosis) was the secondary outcome. Twelve studies totalling 131,822 screened women included eight with European data. All studies measuring test accuracy of AI in screening practice were either retrospective or enriched laboratory test set studies. Low methodological quality according to the QUADAS-2 tool related to concerns about risk of bias and applicability to the clinical context of included studies.

In a retrospective evaluation including 79,910 women, 34/36 (94%) AI systems were less accurate than a single radiologist's original decision; all were less accurate than consensus of two or more radiologists. Five smaller studies (1086 women, 520 cancers) at high risk of bias and low applicability evaluated AI systems as more accurate than a single radiologist reading a test set. In three studies, AI used for triage screened out 53%, 45%, and 50% of women at low risk but also 10%, 4%, and 0% of cancers detected by radiologists.

2.1 Summary of the evidence base [table]

2.1.1 Biomarkers

Technologies	Trial- (Cancer type)	Trial Details	Participants	Outcomes/Results	Notes
Biomarker-protein **(Sun et al., 2020)**	**Serum IDH1 level for early diagnosis of NSCLC**	China NR 1) Training cohort (620) 2) Validation cohort (546)	N = 1223 Mean age NR (17-86 y) 48.9% Male Follow-up NR **Population:** selected subjects with NSCLC or benign pulmonary conditions (BPCs), or other cancers (OC), or good health (healthy control, HC)	**Uptake:** NR **Compliance:** NR **Outcomes:** - **Cancer incidence vs IDH1 level:** In general, the serum IDH1 levels were higher in the NSCLC patients (6.13 ± 4.80 ng/ml) compared to other participants (BPC+OC+HC, 1.90 ± 2.81 ng/ml; $P < 0.001$). When compared specifically with patients with other cancers, the IDH1 level was still higher in NSCLC patients (2.29 ± 3.71 ng/ml vs 6.13 ± 4.80 ng/ml; $P < 0.001$). No difference was observed between healthy control and patients with other cancers in terms of IDH1 serum levels. - **Detection rate:** NR - **Stage:** Using the IDH1 cut off at 5 ng/ml, the specificity for discriminating between early-stage (Stage 0-IA) NSCLC patients and other participants (BPC+OC+HC) in the training cohort was 86.8% with sensitivity of 58.6%. The	Power calculation: NR The serum level of IDH1 was determined using ELISA. The model used for training was not specified.

				PPV was estimated 79.3% and NPV was 70.9%. Likewise in the validating cohort, the specificity was 86.3% with sensitivity of 59.1% for discriminating between early-stage NSCLC patients and other participants (BPC+OC+HC). The corresponding PPV was 53.4% and NPV 88.8%. - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** When the IDH1 cut off was set at 5 ng/ml, the specificity was 92.9% with sensitivity of 63.3% for discriminating between NSCLC patients and healthy subjects in the training cohort. The PPV was estimated 95.4% and NPV was 52.0% Using the same settings, the specificity was 89.3% with sensitivity of 55.0% for discriminating between NSCLC patients and healthy subjects in the validating cohort. The corresponding PPV was 92.7% and NPV 44.6%. - **Model performance:** The AUC of IDH1 values was 0.915 and 0.730 for NSCLC diagnosis in the training and validation cohorts, respectively. When it came to disease stage, the AUC was 0.859 and 0.797 for diagnosing early-stage NSCLC in the training and validation cohorts, respectively.	
Biomarker-protein **(Kazarian et al., 2017)**	**Serum levels of CA15-3, HSP90A and PAI-1 as early diagnosis/pr**	UK 2001-2014 1) UKCTOCS participants who	N = 478 Mean age 61 (50-76 y) Median to diagnosis 13.8 m (up to 5 y)	**Uptake:** NR **Compliance:** NR **Outcomes:**	Power calculation: NR The serum level of all biomarkers was measured using ELISA.

	ognostic markers for BC	developed breast cancer (239) 2) Matched cancer-free control (239)	**Population:** post-menopausal women with BC or matched controls in the UKCTOCS trial	- **Cancer incidence vs biomarker levels:** No biomarker candidates, either alone or in combination, were accurate markers for BC prediction. - **Detection rate:** NR - **Stage:** In relation to clinic-pathological predictions, CA15-3 levels were found raised in samples from late-stage (Stage 3/4) BC patients compared to cancer-free control (*P* = 0.0215). Yet CA15-3 levels were lower in grade 1 BC cases than control (*P* = 0.0254). Serum levels of PAI-1 were significantly lower in patients diagnosed with grade 3 cancer compared to control (*P* = 0.0491). Likewise, HSP90A levels were lower in grade 3 BC cases than cancer-free control (*P* = 0.0174). - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** NR - **Model performance:** The logistic regression model combining all markers for NPI-based prognosis estimated an AUC of 0.77, considering samples taken within 1.15 y of diagnosis.	This was a nested study within the UKCTOCS trial.
Biomarker-protein **(Fitzgerald, di Pietro, O'Donovan, Maroni, et al., 2020; Fitzgerald, Di**	**BEST3** **Cytosponge® combined with TFF3 for early diagnosis of BE, the pre-**	UK 2017-2019 1) Usual care control group (6531) 2) Usual care + offer of	N = 13,514 I = 6983 C = 6531 S = 1654 Median age 69 y 48% Male (among participants taking Cytosponge®)	**Uptake:** Among participants randomised to the intervention group, 39% (2679 of 6983) expressed interest of taking the Cytosponge®-TFF3 procedure, 65% (1750 of 2679) of which met the eligibility criteria and received the procedure **Compliance:** 95% (1654 of 1750) eligible participants successfully swallowed the Cytosponge® for sample production (overall uptake 24%).	Power calculation: Y An endoscopy was offered when TFF3-positive cells were identified in the intervention group or upon advised by general

Pietro, O'Donovan, Muldrew, et al., 2020; Swart et al., 2021)	**cancerous lesion of EAC**	Cytosponge®-TFF3 procedure (6983)	Mean follow-up 12 m **Population:** Patients with long-term symptoms of gastro-oesophageal reflux and received treatment for > 6 m	**Outcomes:** - **Cancer/BE incidence:** Within the follow-up period, 2% (140 of 6834) participants in the intervention group vs <1% (13 of 6388) in the control group were diagnosed with BE (RR 10.6, 95%CI 6.0-18.8; *P* < 0.0001). Nine cases with early-stage neoplasia were diagnosed in the intervention group compared to none in the control group. - **Detection rate:** Among participants taking the Cytosponge®-TFF3 procedure, 13% (221 of 1654) underwent endoscopy due to positive TTF3 results, 59% of which (131 of 221) were diagnosed with BE or EAC. - **Stage:** Among 9 cases of neoplasia diagnosed in the intervention group, 4 were dysplastic BE while 5 were stage-I oesophago-gastric cancer. - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** The specificity of the Cytosponge®-TFF3 procedure for detection of BE, dysplasia, or cancer was estimated 94%. - **Cost-effectiveness:** Compared with usual care, the one-off Cytosponge®-TFF3 screening together with incurred treatment and palliative care for identified BE/EAC led to an incremental of 82 per patient with gastro-oesophageal reflux. An additional 0.015 QALYs per patients was generated with the Cytosponge®-TFF3 screening, rendering an ICER of £5500 per QALY gained. The probabilistic sensitivity analysis revealed an incremental cost of £78 and 0.015 QALYs for Cytosponge®-TFF3	practitioners in the control group.

				screening compared to usual care, giving an ICER of £5405 (95%CI -6791 to 17,600). Considering the willingness-to-pay threshold at £20,000 per QALY, there was a high probability of Cytosponge®-TFF3 being cost-effective than usual care (97%). The total budget impact was also evaluated using the additional cost-per-patient of £82 for one round of Cytosponge®-TFF3 screening, which would cost a total of £21,636,235 spreading over 29 years at an annual cost of £746,077 in the UK settings.	
Biomarker-antibody **(Sullivan et al., 2021)**	**ECLS** **Early CDT-Lung test for predicting LC risk**	UK 2013-2016 1) Usual care control group (6121) 2) EarlyCDT-Lung test + LDCT 6-monthly if test-positive (6087)	N = 12,208 I = 6087 C = 6121 S = NR Mean age 60.5 (50-75 y) 51% Male Mean follow-up 24 m **Population:** former or current smokers [b] or smokers [a] with immediate family history of LC	**Uptake:** NR Compliance: Over 2-year follow-up, the adherence rate to protocol was 89.9%. **Outcomes:** - **Cancer incidence:** The incidence of LC was 520 per 100,000 per annum (0.52%). A total of 56 cases of LC were confirmed in the intervention group (0.92%) whilst 71 cases in the control group (1.16%) within 2 years. - **Detection rate:** Among intervention participants, 9.8% (598 of 6087) were tested positive with EarlyCDT-Lung test and 3.0% (18 of 598) were diagnosed with LC. For those who tested negative, 38 were diagnosed with LC (0.7%). On average, LC patients were diagnosed 87.3 days earlier in the intervention group than control group. - **Stage:** The LC cases of stage III/IV/unspecified were 0.5% (33 of 6087) in the intervention group compared to 0.8% (52 of 6121) in the control group, resulting in a HR of 0.64	Power calculation: Y EarlyCDT-Lung test is a ELISA-based assay, measuring levels of seven autoantibodies in blood samples. The autoantibodies tested include p53, NY-ESO-1, CAGE, GBU4-5, HuD, MAGE A4 and SOX2.

				(95%CI 0.41-0.99; *P* = 0.0432). An estimation of 325 patients was to be screened to prevent one LC case of stage III/IV/unspecified. - **Cancer and all-cause mortality:** No statistically significant difference was observed in terms of LC mortality (0.39% vs 0.28%) or all-cause mortality (1.76% vs 1.43%) between control and intervention arm. - **Sensitivity/Specificity:** The sensitivity of Early CDT-Lung test was 52.2% for detecting stage I/II disease and 18.2% for detecting stage III/IV disease. The corresponding specificity was 90.3% and 90.2%, respectively. The PPV was estimated 2.0% for stage I/II disease and 1.0% for stage III/IV disease while the corresponding NPV was 99.8% for the former and 99.5% for the latter. - **Cost-effectiveness:** The cost per stage I/II LC detected after 2 years was £116,000.	
Biomarker-protein **(Tarney et al., 2019)**	**Biomarker panel for early detection of endometrial cancer**	US 1993-2001 1) Endometrial cancer cases (112) 2) Cancer-free matched control (NR)	N = NR Mean age NR Median follow-up 17 y (as PLCO) **Population:** postmenopausal women with endometrial cancer or matched controls in PLCO trial	**Uptake:** NR **Compliance:** NR **Outcomes:** - **Cancer incidence:** NR - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** Forty seven proteins were found abundant differentially between cancer cases and matched controls (*P* < 0.05). The integrated risk score of	Power calculation: NR This was a nested study in the PLCO trial, where cases of endometrial cancer were matched with control for quantitative proteomics and phosphoproteomics of pre-diagnostic serum.

				6 proteins including complement factor B, serotransferrin, catalase, proteasome subunit beta type-6, beta-2-microglobulin, and protocadherin-18 were found directly associated with cancer incidence. - **Model performance:** The AUC of the integrated biomarker panel for distinguishing cancer case and control was 0.80 (95%CI 0.72-0.88).	
Biomarker-protein **(In et al., 2021)**	**Serum pepsinogen as a biomarker for GC**	US 1993-2001 1) Gastric cancer cases (105: 70 non-cardia and 35 cardia) 2) Cancer-free matched controls (220)	N = 325 Mean age NR Median follow-up 17 y (as PLCO) **Population:** GC cases and matched controls in PLCO trial	**Uptake:** NR **Compliance:** NR **Outcomes:** - **Cancer incidence:** NR - **Detection rate:** Higher PG+ rate was observed in GC patients compared to controls (31.4% vs 5.5%; *P* < 0.001). The risk of GC was significantly higher in PG+ than PG- participants, leading to a HR of 3.77 (95%CI 2.50-5.71; adjusted HR 4.42; 95%CI 3.14-6.21). Among sub-cohort of non-cardia GC, PG+ demonstrated an increased risk of GC compared to PG- (HR 5.65; 95%CI 3.67-8.70; adjusted HR 7.26; 95%CI 4.84-10.90). Yet such trend was not found in the cardia GC sub-cohort (HR 1.79; 95%CI 0.72-4.44; adjusted HR 1.95; 95%CI 0.81-5.37). - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** NR	Power calculation: NR This was a nested study in the PLCO trial, where serum samples of GC cases were compared with those of control in terms of PG level using ELISA. HR was adjusted for family history of GC, smoking and BMI.

Biomarker - protein **(Chiamulera et al., 2021)**	**Systematic review - Salivary cytokines as biomarkers for oral cancer**	Multiple countries 2004-2018 28 case-control studies included	N = 18-300 Mean age NR Follow-up NR **Population:** not specified	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs salivary cytokine:** Compared with healthy controls, levels of specific salivary cytokines were significantly different in oral cancer patients: IL-8 (SMD 1.77; 95%CI 0.79-1.55), IL-6 (SMD 2.08; 95%CI 1.33-2.84), TNF-a (SMD 2.04; 95%CI 0.47-3.61), IL-1b (SMD 0.78 95%CI 0.44-1.13), IL-10 (SMD 0.46; 95%CI 0.05-0.86), IL-1a (SMD 2.21; 95%CI -0.36-4.77). - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** NR	Power calculation: NR Only studies using ELISA for measuring salivary cytokines were included. The frequency of salivary cytokines examined were IL-8 (50%), IL-6 (50%), TNF-a (28.6%), IL-1b (21.4%), IL-10 (17.9%), IL-1a (10.7%), and IL-1, IL-1RA, IL-4 and IL-13 (3.6% each).
Biomarker - protein **(Aalami et al., 2021)**	**Systematic review – Urinary angiogenin as biomarkers for bladder cancer**	Egypt/USA 2004-2014 4 case-control studies included	N = 656* Mean age NR Follow-up NR **Population:** not specified *Pooled from all 4 RCTs	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs urinary angiogenin:** NR - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR	Power calculation: NR All included studies used ELISA for measuring urinary angiogenin levels despite varied cut-off.

				- **Sensitivity/Specificity:** The analysis of pooled studies revealed a sensitivity of 0.71 (95%CI 0.60-0.75), specificity of 0.78 (95%CI 0.73-0.81), positive likelihood ratio of 3.34 (95%CI 2.02-5.53), negative likelihood ratio of 0.37 (95%CI 0.32-0.44), diagnostic odds ratio of 9.99 (95%CI 4.69-21.28) and AUC of 0.789.	
Biomarker-DNA **(F. Carozzi et al., 2017; F. M. Carozzi et al., 2017)**	**ITALUNG Biomarker Panel (IBP) for LC detection**	Italy 2004-2006 1) Lung cancer cases (36) 2) Cancer-free matched controls (481)	N = 517 Mean age 61.1 y 60.2% Male Follow-up 4 y **Population:** LC cases and matched controls in ITALUNG trial	**Uptake:** NR **Compliance:** Of 1,406 screened, 1,356 (96%) consented to give a sample of blood and sputum at baseline. Random selection then made of 517 samples. **Outcomes:** - **Cancer incidence:** Among 36 LC cases enrolled, 18 were detected at the baseline LDCT screening and another 18 detected at annual repeat LDCT. - **Detection rate:** Among 517 subjected screened, 146 were LDCT positive. The IBP positive rate among LC cases was 94.4% (17 of 18) and 66.7% (12 of 18) at baseline and repeat screening, respectively. - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** Based on the data of this cohort, the specificity of LDCT alone was 74% while IBP alone was 59%. The combination of both led to an improved specificity of 90% and PPV of 26%. A simulation was carried out to extrapolate the findings and found the sensitivity being the same (90%) for either approach	Power calculation: NR ITALUNG biomarker study was aimed to evaluate the efficiency of combining molecular markers and LDCT as a screening approach. Plasma DNA was quantified with real-time PCR while blood and sputum samples were subjected to assessment of MSI and LOH.

				alone at baseline with lower specificity for IBP (61% vs 71% for LDCT). The PPV was comparable, 4.3% for LDCT vs 3.3% for IBP. In terms of multimodal approach where LDCT was combined with IBP for screening, the specificity was improved to 89% as well as the PPV (10.6%) with unchanged sensitivity (90%). The probability of LC confirmation under circumstances of LDCT negative and IBP positive was estimated 3.4% throughout the whole screening cycle.	
Biomarker - DNA methylation **(Zhou et al., 2020)**	**Systematic review – Methylation status of the APC promoter and GC risk**	Multiple countries 2003-2015 8 case-control studies included	N = 985* Mean age NR Follow-up NR **Population:** not specified *Pooled from all 8 RCTs	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs methylation of the APC promoter:** Higher methylation of APC promoter was observed in patients with GC compared to patients without GC (OR 3.86; 95%CI 1.71-8.74; *P* = 0.001). - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** NR	Power calculation: NR Tissues or blood specimen from patients were used for assessing the methylation status of APC promoter.
Biomarker - DNA methylation	**DNA hypermethylation of biomarkers for LC**	The Netherlands and Belgium 2003-2006	N = 284 Mean age NR % Male NR	**Uptake:** NR **Compliance:** Sputum was collected from 1,548 (20%) of 7,915 subjects in the LDCT screening arm. Samples then identified for analysis.	Power calculation: NR The DNA hypermethylation of the following biomarkers in

(Hubers et al., 2017)	**detection in the NELSON trial**	1) Lung cancer cases (65) 2) Cancer-free controls with minor cytological aberrations (120) 3) Cancer-free controls without cytological aberrations (99)	Follow-up 80 m **Population:** LC cases and matched controls in NELSON trial	**Outcomes:** - **Cancer incidence:** NR - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** The DNA hypermethylation of RASSF1A may be useful for detecting invasive LC in a screening interval of 2 years with specificity of 93% (95%CI 89-96%) and sensitivity of 17% (95%CI 4-31%). Within 2-year interval, the biomarker panel consisting of RASSF1A, 3OST2 and PRDM14 could detect 28% of LC cases with specificity of 90% (95%CI 86-94%).	the sputum samples were examined: RASSF1A, APC, cytoglobin, 3OST2, FAM19A4, PHACTR3 and PRDM14. The cut-off values were determined for high specificity of diagnostic value assessment per biomarker.
Biomarker - DNA methylation **(Sturgeon et al., 2021)**	**DNA methylation in WBC as biomarker for BC - CpG sites**	US 1993-2001 1) Breast cancer cases (297) 2) Cancer-free matched controls (297)	N = 594 Mean age NR (55-74 y) Follow-up 17 y (as PLCO) **Population:** BC cases and matched controls in the intervention arm of PLCO trial	**Uptake:** NR **Compliance:** Overall 97% of participants provided two serial WBC DNA samples for analysis. On average, proximate samples were taken 1.82 years before diagnosis whilst distant samples were taken 5.7 years prior to diagnosis. **Outcomes:** - **Cancer incidence vs CpG sites:** One percentage increase in *ERCC1* CpG site in proximate WBC DNA samples was associated with increased BC risk (adjusted OR 1.29; 95%CI 1.06-1.57), but an inversely association was	Power calculation: NR This was a nested study in the PLCO trial, where blood samples of BC cases were compared with those of control in terms of CpG sites using targeted bisulphite amplification sequencing.

				observed in distant WBC DNA samples (adjusted OR 0.83; 95%CI 0.69-0.98). - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** NR	
Biomarker - DNA methylation **(Sturgeon et al., 2017)**	**DNA methylation in WBC as biomarker for BC - %5-mdC**	US 1997-2005 1) Invasive breast cancer cases (428) 2) Cancer-free matched controls (419)	N = 847 Mean age NR (55-74 y) Follow-up 17 y (as PLCO) **Population:** BC cases and matched controls in the intervention arm of PLCO trial	**Uptake:** NR **Compliance:** NR **Outcomes:** - **Cancer incidence vs %5mdc levels:** No correlation was observed between DNA methylation in WBC samples and breast cancer risk. - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** NR	Power calculation: NR This was a nested study in the PLCO trial, (Etiology and Early Marker Study, EEMS) where blood samples of BC cases were compared with those of control in terms of ratio of 5-mdC to dG using liquid chromatography-electrospray ionization-tandem mass spectrometry.
Biomarker - DNA methylation **(Guo et al., 2019)**	**Systematic review - Promoter methylation as biomarkers for**	Multiple countries NR	N = 7-206 Mean age NR Follow-up NR **Population:** not specified	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs promoter methylation:** Among 50 genes reported across 33 studies, hypermethylation of	Power calculation: NR In total the promoter methylation of 50 genes were reported in studies included.

	melanoma diagnosis	33 case-control studies included		the following genes were found higher in melanoma patients than in cancer-free controls: CLDN11 (OR 16.82; 95%CI 1.97-143.29; *P* = 0.010, MGMT (OR 5.59; 95%CI 2.51-12.47; *P* < 0.0001), p16 (OR 6.57; 95%CI 2.19-19.75; *P* = 0.0008), RAR-b2 (OR 24.31; 95%CI 4.58-129.01; *P* = 0.0002) and RASSF1A (OR 9.35; 95%CI 4.73-18.45; *P* < 0.00001). - **Detection rate:** NR - **Stage:** In terms of disease stage, hypermethylation of CLDN11 (OR 14.52; 95%CI 1.84-114.55; *P* = 0.01), MGMT (OR 8.08; 95%CI 1.84-35.46; *P* = 0.006), p16 (OR 9.44; 95%CI 2.68-33.29; *P* = 0.0005) and RASSF1A (OR 7.72; 95%CI 1.05-56.50; *P* = 0.04) were found increased in primary melanoma compared with controls. When it comes to metastasis melanoma, the methylation frequency of CLDN11 (OR 25.56; 95%CI 2.32-281.66; *P* = 0.008), MGMT (OR 4.64; 95%CI 1.98-10.90; *P* = 0.0004), p16 (OR 4.31; 95%CI 1.33-13.96; *P* = 0.01) and RASSF1A (OR 10.10; 95%CI 2.87-35.54; *P* = 0.0003) was significantly higher in patients compared with controls. - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** NR	
Biomarker - DNA methylation **(Anghel et al., 2021)**	**Systematic review - Promoter methylation as biomarkers**	Multiple countries NR 74 diagnostic accuracy	N = NR Mean age NR Follow-up NR	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs promoter methylation:** NR	Power calculation: NR Currently approved epigenetic tests for CRC screening are: ColoGuard® (US), Epi

	for CRC early detection	studies included	**Population:** not specified	- **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** DNA methylation of SDC2 was estimated to have a sensitivity of 83.3-91.3% for detecting stage I/II disease and 89.6-100% for detecting stage III/IV disease. The SEPT9 methylation assessment processed 100% sensitivity for stage I disease when use in combination with FOBT. A gene panel capable of testing methylation of SDC2 and SEPT9 demonstrated a sensitivity of 69.1-81.8% for stage I disease, 85.7%-100% for stage II disease, 88.9-89.7% for stage III disease and 75-100% for stage IV disease. Though not yet approved, another panel testing 5 CTCF binding sites showed 93.54% sensitivity and 94.05% specificity for CRC detection.	proColon® (US), EarlyTect®-Colon Cancer (Korea) and Colosafe® (China). ColoGuard® is the first FDA-approved stool CRC screening kit, testing methylation level of NDRG4 and BMP3 as well as mutations of KRAS and □-actin. Epi proColon is the first FDA-approved blood-based screening kit, testing the methylation of SEPT9. SDC2 and SEPT9 were the most frequently assessed epigenetic markers for CRC detection
Biomarker - extracellular RNA (miRNA, lncRNA, or circular RNA)	**Systematic review - RNA biomarkers in biofluids for early diagnosis of**	Multiple countries NR	N = 50-3079 Mean age NR Follow-up NR **Population:** not specified	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:**	Power calculation: NR The studies included were mostly focusing on blood-derived fluids (34 of 26). Only one study looked into urine while

(Hulstaert et al., 2021)	**ovarian cancer**	36 case-control studies included		- **Cancer incidence vs RNA biomarkers:** Higher levels of miR-21, the miR-200 family, miR-205, miR-10a and miR-346 were observed in biofluids of cancer patients compared to controls. In contrast, levels of miR-122, miR-193a, miR-223, miR-126 and miR-106b were lower. - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** NR - **Model performance:** The best RNA biomarkers reported had AUCs ranged between 0.694 to 1. The best performing model with validation was a panel consisting of 10 miRNAs (miR-320a, miR-665, miR-3184-5p, miR-6717-5p, miR-4459, miR-6076, miR-3195, miR-1275, miR-3185 and miR-4640-5p), with an AUC of 1, sensitivity of 0.99 and specificity of 1. The second-best performing model consisted of 4 miRNAs (miR-7, miR-429, miR-25 and miR-93) with an AUC of 0.98, sensitivity of 0.93 and specificity of 0.92.	another checked ascites. The method categories across included studies were reverse transcription quantitative PCR, microarray and RNA-sequencing.
Biomarker - extracellular RNA (miRNA, lncRNA, or circular RNA) **(Yu et al., 2020)**	**Systematic review - lncRNA as biomarkers for early diagnosis of digestive tract cancer**	Multiple countries NR 69 diagnostic accuracy studies included (40 in	N = NR Mean age NR Follow-up NR **Population:** not specified	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs RNA biomarkers:** NR - **Detection rate:** NR - **Stage:** NR	Power calculation: NR The specimens included blood, tissue and, for GC, also gastric juice.

		GC; 24 in CRC; 5 in EC)		- **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** In general, the sensitivity and specificity of lncRNA in screening digestive track cancer was 0.78 and 0.80, respectively. The corresponding rate for cancer type was: 0.77 (95%CI 0.72-0.81) and 0.75 (95%CI 0.71-0.79) for GC; 0.82 (95%CI 0.76-0.86) and 0.84 (95%CI 0.79-0.88) for CRC; 0.74 (95%CI 0.67-0.80) and 0.86 (95%CI 0.72-0.93) for EC. - **Model performance:** The overall AUC of lncRNA in screening digestive track cancer was 0.86. In terms of each cancer type, the AUC was 0.83 for GC; 0.90 for CRC; 0.82 for EC.	
Biomarker - extracellular RNA (miRNA, lncRNA, or circular RNA) **(Chu et al., 2018)**	**Systematic review - miRNA signature classifier (MSC) as biomarkers for LC detection in MILD trial**	Italy 2005-2011 939 plasma samples (69 from LC patients)	N = 939 Mean age NR (≥ 50 y) Follow-up 5 y **Population:** former [d] or current [b] smokers	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs RNA biomarkers:** NR - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** A total of 19 patients died due to LC and no participants died because of other cause during the follow-up. For LC mortality, MSC demonstrated a sensitivity of 95% (95%CI NR), specificity of 78% (95%CI 75-80%), PPV of 8% (95%CI 5-12%) and NPV of 99% (95%CI NR). - **Sensitivity/Specificity:** For LC detection, MSC demonstrated a sensitivity of 87% (95%CI NR), specificity	Power calculation: NR MSC is a plasma-based miRNA panel including 24 miRNAs, based on which patients are categorized into low, intermediate, or high risk of LC.

				of 81% (95%CI 79-84%), PPV of 27% (95%CI 21-32%) and NPV of 98% (95%CI NR).	
Biomarker - extracellular RNA (miRNA, lncRNA, or circular RNA) **(Chu et al., 2018)**	**Systematic review - miR-test as biomarkers for LC detection in COSMOS trial**	Italy 2004-2005 1008 serum samples (36 from LC patients)	N = 1008 Mean age NR (> 50 y) Follow-up NR **Population:** Current [b] or former smokers	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs RNA biomarkers:** NR - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** Three patients died of LC during the follow-up. No all-cause mortality analysis was provided. For LC mortality, miR-test demonstrated a sensitivity of 100% (95%CI NR), specificity of 73% (95%CI 70-76%), PPV of 1.1% (95%CI NR) and NPV of 100% (95%CI NR). - **Sensitivity/Specificity:** For LC detection, miR-test demonstrated a sensitivity of 78% (95%CI NR), specificity of 75% (95%CI 72-78%), PPV of 10% (95%CI 7-14%) and NPV of 98% (95%CI NR).	Power calculation: NR MiR-test is a serum-based miRNA panel including 13 miRNAs, including miR-92a-3p, miR-30b-5p, miR-191-5p, miR-484, miR-328-3p, miR-30c-5p, miR-374-5p, let-7d-5p, miR-331-3p, miR-29a-3p, miR-148a-3p, miR-223-3p and miR-140-5p.
Biomarker - extracellular RNA (miRNA, lncRNA, or circular RNA)	**Systematic review - miRNA-21 as biomarkers for detecting colorectal**	Multiple countries NR 11 studies included	N = 2139* Mean age NR Follow-up NR **Population:** not specified	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs RNA biomarkers:** NR - **Detection rate:** NR	Power calculation: NR The specimens included tissue (12 out of 14 cohorts) and serum.

(Saheb Sharif-Askari et al., 2020)	**adenocarcinoma**		*Pooled from all 11 RCTs	- **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:**.NR - **Survival:** The meta-analysis revealed that high level of miR-21 was associated with worse overall survival (HR 1.75; 95%CI 1.23-2.51; *P* = 0.001). Despite a trend between miR-21 overexpression and disease-free survival, it was not statistically significant (HR 1.21; 95%CI 0.91-1.60; *P* = 0.19).	
Biomarker - liquid biopsies **(Campi et al., 2021)**	**Systematic review – Novel liquid biomarkers and renal cell carcinoma**	Multiple countries 2016-2019 6 case-control studies included	N = NA Mean age NR Follow-up NR **Population:** not specified	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs liquid biomarkers:** Inconclusive - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** NR	Power calculation: NR Liquid biopsies including urine, plasma and serum was utilised for assessing the extracellular RNAs or metabolites.
Others **(Yang et al., 2019)**	**DNA quantitative cytology for detection of endometrial cancer**	China 2013-2017 1) Non-menopausal women (NR)	N = 575 Mean age NR Follow-up NR **Population:** general female population	**Uptake:** NR **Compliance:** NR **Outcomes:** - **Cancer/pre-cancerous lesion incidence:** Among 575 women enrolled, 47 endometrial cancer cases were	Power calculation: NR All participants went through endometrial DNA quantitative cytology tests and hysteroscope plus dilation and curettage.

		2) Menopausal women (NR)		confirmed, 30 were diagnosed with atypical hyperplasia, 382 were with benign lesion and 116 were normal. - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** The accuracy of DNA quantitative cytology for diagnosing endometrial cancer was 85.57%, with a sensitivity of 87.01%, specificity of 85.34%, false negative rate of 12.99%, false positive rate of 14.66%, PPV of 47.86% and NPV of 97.07%. In terms of detection in menopausal women, the accuracy of DNA quantitative cytology was 89.95%, with a sensitivity of 97.73%, specificity of 87.59%, false negative rate of 2.27%, false positive rate of 12.41%, PPV of 70.49% and NPV of 99.22%. For detection in non-menopausal women, the accuracy of DNA quantitative cytology was 83.42%, with a sensitivity of 72.73%, specificity of 84.42%, false negative rate of 27.27%, false positive rate of 15.58%, PPV of 30.38% and NPV of 97.07%.	This study was aimed to compare the efficiency of DNA quantitative analysis with the clinical histopathological results in terms of endometrial cancer detection.
Others (Wen et al., 2021)	**Systematic review – Urinary volatile organic compound (VOC) analysis for**	Multiple countries 1999-2019 13 case-control studies across 5	N = 1266* of which 700 were diagnosed with cancer Mean age NR Follow-up NR **Population:** not specified	**Uptake:** Not applicable **Compliance:** Not applicable **Outcomes:** - **Cancer incidence vs VOC contents:** In total, 48 urinary VOCs belonging to 11 chemical classes were found associated with cancers. Twenty-nine urinary VOCs were identified for PC, most of which decreased in the urine of	Power calculation: NR Across 13 studies, 10 studies analysed the urinary samples with gas chromatography mass spectrometry (GC-MS); 3 used selected ion flow tube mass

	cancer diagnosis	cancer types included	*Pooled from all 13 RCTs	cancer patients compared to non-cancer patients. Distinct set of VOCs were identified for GCs with 19 out the 21 cancer-associated VOCs different from those of PC, most of which increased in cancer patient compared to non-cancer patients. For leukaemia/lymphoma, 6 VOCs were found mostly increased in the urine of patients except for anisole. In the case of bladder cancer, formaldehydes were reported as VOCs associated with the malignancy whilst no urinary VOC was found associated specifically with LC. - **Detection rate:** NR - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** Within included studies, VOCs associated with PC and GCs demonstrated high sensitivity and specificity for cancer detection.	spectrometry (SIFT-MS); a single study used field asymmetric ion mobility spectrometry (FAIMS). Nine out of 13 studies analysed VOCs within the headspace instead of the fluid phase of urine.

[a] < 20 pack-y
[b] ≥ 20 pack-y
[c] > 2 h-day for at least 10 y
[d] ≤ 10 y since quitting
[e] Quit after age 50 and < 10 y since quitting
[f] ≥ 20 pack-y in the last 10 y or quit < 10 y
[g] ≥ 30 pack-y
[h] 15 cigarettes/d for ≥ 25 y or 10 cigarettes/d for ≥ 30 y
[i] 15 cigarettes/d for >25 y or 10 cigarettes/d for > 30 y
[j] ≥ 15 cigarettes/d for ≥ 20 y

Abbreviations and acronyms	
	Uptake: Percentage of invited population agreeing to participate in the trial *Compliance:* Percentage of trial population providing samples for analysis N=Total number in trial; I=in intervention group(s); C= in control group; S=No. screened
5-mdC	5-methyl-2' deoxycytidine
AI	Artificial intelligence
APC	Adenomatous polyposis coli
AUC	Area under the curve
BC	Breast cancer
BE	Barrett's oesophagus
BEST3	Barrett's OESophagus Trial 3
BMI	Body mass index
BMP3	Bone morphogenic protein 3
CAI	Colonoscopy with air method
CC	Cervical cancer
CIN	Cervical intraepithelial neoplasia
CLC	Colorectal cancer
CLDN11	Claudin 11
COSMOS	Continuous Observation of Smoking Subjects trial
CpG	Cytosines followed by guanine on the same strand of DNA and connected by a phosphate
CRC	Colorectal cancer
CTCF	CCCTC-binding factor
CWE	Chromoendoscopy & water exchange
dG	2'-deoxyguanine
DTS	Chest digital tomosynthesis
EAC	Oesophageal adenocarcinoma
EC	Oesophageal cancer
ECLS	Early Diagnosis of Lung Cancer Scotland
ELISA	Enzyme-linked immunosorbent assay
GC	Gastric cancer
GI	Gastrointestinal cancer

HR	Hazard ratio
HPV	Human papillomavirus
IBP	ITALUNG Biomarker Panel
ICER	Incremental cost-effectiveness ratio
IDH1	Isocitrate dehydrogenase 1
IT	Information technology
LBC	Liquid-based cytology
LC	Lung cancer
lncRNA	Long non-coding RNA
LOH	Loss of heterozygosity
Lung-RADS	Lung imaging reporting and data system
LYG	Life-years gained
m	Month
MGMT	O-6-methylguanine-DNA methyltransferase
miRNA	microRNA
MSC	miRNA signature classifier
MSI	Microsatellite instability
NA	Not applicable
NDRG4	N-Myc downstream-regulated gene 4 protein
NPI	Nottingham prognostic index
NR	Not reported
NSCLC	Non-small-cell lung cancer
OR	Odds ratio
Pap-smear	Papanicolaou cytology
PC	Prostate cancer
PG	Pepsinogen
PLCO	Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial
PPV	Positive predictive value

PSA	Prostate specific antigen
QALY	Quality adjusted life years
RAR-b2	Retinoic acid receptor b
RASSF1A	Ras association domain family member
RCT	Randomised controlled trial
ROC	Receiver operating characteristic
RR	Risk ratio
SDC2	Syndecan 2
SEPT9	Septin 9
SMD	Standardised mean difference
SOS	Studio OSservazionale
TFF3	Trefoil factor 3
TRIPOD	Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis
VOC	Volatile organic compound
WBC	White blood cells
WE	Water exchange colonoscopy
y	Year

2.1.2 Imaging and artificial intelligence

Technologies	Trial citation	Trial Details	Participants	Outcomes/Results	Notes
AI & imaging	Freeman et al. 2021	Systematic review of detection accuracy of standalone AI	**Total number (no. screened):** N = 131,822	**Uptake:** Not applicable **Compliance:** Not applicable	

		algorithms or AI-assisted radiologists in ca breast mammography UK study including studies from multiple countries 2010–2021 12 retrospective studies Endpoints: 1. Test accuracy 2. Ca breast type detected	**Population:** women screened within digital mammography programmes	**Outcomes:** **Sensitivity/Specificity:** For larger retrospective studies (pooled n = 79,910), specificity of standalone AI systems was lower for 94% (34/36) systems vs. single radiologist detection and 100% for two radiologist consensus. Smaller laboratory studies (pooled n =1086) reported AI to be more accurate than a single radiologist. **Harm-benefit:** AI used to triage for radiological review screened out 45–53% of women at low risk but also up to 10% of radiologist detected cancers **Incidence/Stage:** One AI system detected fewer cases of DCIS than radiologists (83.5% vs. 89.4%) and more invasive BCa (82.8% vs 76.7% & 79.7%) and more ≥ Stage 2 cancers (78.4% vs. 68.1%). Two other AI systems detected less ≥ Stage 2 cancers Mortality: NR **Cost-effectiveness:** NR	
Imaging **One-time 2nd-generation NBI vs WLI** **Gastric cancer**	Yoshida et al. 2021	Open-label crossover RCT Japan 2014–2017 Endpoints:	**Total number (no. screened):** 4575 (4472) **Population:** High-risk of GC*	**Uptake:** NR **Compliance:** n = 2234/2258 (99%) in primary WLI group and n = 2238/2265(99%) in primary NBI group **Outcomes:**	Power calculation: Y [revised after interim analysis] *High risk defined as 20–85 years with either a:

		1. Detection rate of EGC 2. PPV; observation time; missed ECG in primary exam		**Incidence/Stage:** EGC incidence rate was 44 (1.9%) vs 53 (2.3%) for primary WLI vs NBI. Overall rate of lesions detected at secondary examination was 25% (n=36/145) with no significant difference between groups. Mortality: NR **Harm-benefit:** PPV for EGC in suspicious lesions was 13.7% (50/372) vs 20.9% (59/282), for WLI vs NBI ($P = 0.015$). Mean observation time was 233 sec and 253 sec for WLI and NBI respectively ($P < 0.001$). **Cost-effectiveness:** NR	(1) history of endoscopic resection for an oesophageal cancer or gastric neoplasm; (2) current oesophageal cancer or gastric neoplasm; (3) history of chemotherapy and/or radiation therapy for oesophageal cancer.
Imaging **One-time LCI+WLI vs WLI** **Gastric cancer**	Gao et al. 2021	RCT China Data collection dates: NR Endpoint: Detection of gastric neoplastic lesions	**Total number (no. screened):** 2383 (2335) **Population:** High-risk of EGC*	**Uptake:** NR **Compliance:** 96% (1110/1160) in WLI group and 100% (1125) in LCI+WLI group were observed. **Outcomes:** **Incidence/Stage:** EGC incidence was 4.3% (50/1110) in WLI group vs 8.0% (98/1125) in LCI+WLI group: a detection rate difference = 3.7% (95%CI 1.36–2.75, P <0.001). Detection of type IIb lesions and high-grade precursor lesions was significantly higher in LCI+WLI group (both P =0.01). Mortality: NR **Harm-benefit:** NR **Cost-effectiveness:** NR	Power calculation : Y * Definition of high-risk based on the Consensus on Screening and Endoscopic Diagnosis and Treatment of Early Gastric Cancer in China (2014)

Imaging **One-time BLI-b vs WLI** **Gastric cancer**	Dohi et al. 2019	Crossover RCT Japan 2013-2017 Endpoints: 1. Detection of EGC by primary imaging exam 2. Detection of EGC by secondary imaging exam	**Total number (no. screened):** 629 (596) **Population:** High-risk of EGC (atrophic gastritis with intestinal metaplasia or surveillance after endoscopic resection of EGC)	**Uptake:** NR **Compliance:** 100%; n = 298 in primary WLI group and n = 298 in primary BLI-b group were observed using both imaging technologies. **Outcomes:** **Incidence/Stage:** EGC incidence was 7.0% (21/298) in primary WLI group vs 8.7% (26/298) in primary BLI-b group. The real-time detection rate of primary WLI was 50.0% vs 93.1% for BLI-b. BLI-b had significantly greater detection of smaller/earlier stage GC (<10 mm and 10—20 mm; lesions with depth of invasion of T1a) plus other pathomorphological types (open atrophic border; lesions in lower 1/3 of stomach; flat lesions; well-differentiated adenocarcinomas.) Mortality: NR **Harm-benefit:** NR **Cost-effectiveness:** NR	Power calculation : Y BLI-bright mode = addition of a control for the 2 lasers along with white-light-emitting phosphors

Imaging **One-time WLI+NBI vs WLI+Lugol chromoendos-copy** **Oesophageal cancer**	Gruner et al. 2021	RCT France 2011–2015 Endpoint: 1. Specificity of detection of oesophageal SCC and HGD 2. Sensitivity (PPV, NPV)	**Total number (no. screened):** 334 (316) **Population:** History of SCC of UAD tract and scheduled for gastroscopy	**Uptake:** NR **Compliance:** n= 8 and n = 11 did not receive the allocated Lugol and NBI examination, respectively. Overall compliance of 315/334 = 94%. **Outcomes:** **Incidence/Stage:** 18/106 (17.0%) of suspected lesions detected by Lugol were confirmed as SCC (14), HGD (1) and LGD (3). Lugol detected 7 additional neoplastic lesions after WLI. 22/61 (36.1%) of suspected lesions detected by NBI were confirmed as SCC (20 T1, 2 T2). 21 of these lesions had been detected by WLI. There was no statistically significant difference in number of patients with HG lesions detected between Lugol and NBI groups (8.4% vs. 10.8%; *P* =0.58). Mortality: NR **Harm-benefit:** Specificity was greater with NBI than Lugol (*P* =0.002). In per-patient analysis, sensitivity, specificity, PPV and NPV were 100%, 66.0%, 21.2%, and 100%, respectively for Lugol vs 100%, 79.9%, 37.5%,and 100%, respectively for NBI. **Cost-effectiveness:** NR	Power calculation : Y
Imaging	Ferreira et al. 2021	RCT	**Total number (no. screened):**	**Uptake:** NR **Compliance:** NR	Power calculation : N

One-time BLI-b vs LCI vs WLI **Colorectal cancer**	[conference abstract]	Brazil Endpoint: Detection of CRC adenoma	168 **Population:** average risk of CRC adenoma	**Outcomes:** **Incidence/Stage:** Overall detection rate was 60.1%: 55.5% for WLI; 55.5% for BLI-b; 68.3% for LCI (*P* =0.03). All technologies were similar at detecting lesions <5mm dia, but LCI had superior flat-lesion detection ability. Mortality: NR **Harm-benefit:** NR **Cost-effectiveness:** NR	
Imaging **TD vs standard online CTC** **Skin cancer**	Ferrándiz et al. 2017	RCT Spain 2015 Endpoint: 1. Diagnostic performance 2. Cost-effectiveness	**Total number (no. screened):** 454 (454) **Population:** adults accessing primary care with concerning skin lesions	**Uptake:** NR **Compliance:** 100%; All n = 226 in CTC group and n = 228 in TD group were examined. **Outcomes:** **Incidence/Stage:** Proportion referred for in-person evaluation was 45.1% (95%CI 38.7–51.6) for CTC vs 20.9% (95%CI 15.0–25.4) for TD (P < 0.001). Mortality: NR **Harm-benefit:** Sensitivity and specificity were significantly higher for TD (92.9% & 96.2%) vs CTC (86.6% & 72.3%). The accuracy index was 94.3% for TD and 79.2% for CTC (*P* <0.001): OR of correct diagnosis using TD = 4.04 (95%CI 2.0–8.1; *P* < .0001). **Cost-effectiveness:** TD was the dominant strategy, with a lower cost-effectiveness ratio	Power calculation: NR

				(65.13€ vs 80.84€/accuracy units). The time dedicated by the TD operator in managing teleconsultation was 7.2 mins (95%CI 6.8–7.6) for CTC and 8.9 minutes (95%CI 8.3–9.5) for TD (*P* <0.001).	
Imaging **Initial round of Helical CT screening from NLST** **Lung cancer**	NLST dataset Hostetter et al. 2017	Modelling study of personalised malignancy risk US	**Total number (no. screened):** 53,454 (26,722) **Population:** 55–75 yrs, smoking history ≥30 pack-years in CT arm of NSLT	**Uptake:** NR **Compliance:** Data from n = 26,722 in CT group of NSLT. **Outcomes:** **Incidence/Stage:** 5840 had lung nodules of any size at initial screening, with 465 cancers in same lobe as largest nodules: a prevalence of malignancy in the nodules of 8.5%. Nodule size predicted malignancy risk; prevalence of cancer in nodules ≤4mm was 3.16% vs 21.79% in nodules >8 mm. Additional significant risk stratification discriminators were smoking history, sex, and nodule location. Mortality: NR **Harm-benefit:** Using personalised malignancy risk model, 54% of nodules >4 and ≤6 mm were reclassified to longer-term FU than recommended by non-personalised criteria. 27% of nodules ≤4 mm were reclassified to shorter-term FU **Cost-effectiveness:** NR	Power calculation: Y for underlying NLST

AI in conventional imaging **Lung cancer**	**SOS** (Chauvie et al., 2020)	Italy Non-randomised study 2010-2018 1) Binary visual analysis 2) Lung-RADS classification 3) Logistic regression (LR) 4) Random Forest (RF) 5) Neural network (NNET)	N = 1594 Mean age 63.2 (45-75 y) 65% Male ≥ 1 y follow-up **Population:** former [d] or current smokers [b]	**Uptake:** NR **Compliance:** NR **Outcomes:** - **Cancer incidence:** A total of 32 lung cancer cases were diagnosed, one of which was not identified via DTS. - **Detection rate:** Over 3 rounds of DTS screening, results of 234 participants were positive. - **Stage:** NR - **Cancer and all-cause mortality:** NR - **Sensitivity/Specificity:** The corresponding sensitivity of method 1 to 5 was 95%, 65%, 20%, 30% and 90% with comparable specificity (93-100%). The PPV of binary visual analysis was the lowest (14%) followed by Lung-RADS (19%), LR (29%), and then RF (40%) whilst NNET had highest PPV of 95%.	Power calculation: NR Report was developed in accordance with TRIPOD guidelines. This trial was aimed to evaluate whether AI can enhance the sensitivity and specificity of DTS in lung cancer detection. DTS was performed using Discovery XR650 (GE Healthcare) with tube voltage of 120 kVp. Both semantic variables and radiomics features were used to develop a LG-based prediction model and machine learning.

Acronym	Full Description
	Uptake: Percentage of invited population agreeing to participate in the trial *Compliance:* Percentage of trial population completing the baseline screening
BCa	Breast cancer
BLI-b	Blue laser imaging - bright mode
CRC	Colorectal cancer
CTC	Clinical teleconsultation
DCIS	Ductal carcinoma in situ
DTS	Chest digital tomosynthesis
EGC	Early gastric cancer
HGD	High grade dysplasia
LCI	Linked colour imaging
LGD	Low grade dysplasia
NBI	Narrow band imaging
NLST	National Lung Screening Trial
NPV	Negative predictive value
PPV	Positive predictive value
QUADAS-2	QUality Assessment of Diagnostic Accuracy Studies-2
SCC	Squamous cell carcinoma
TD	Tele-dermoscopy
UAD	Upper aerodigestive (tract)
WLI	White light imaging

2.2 Bottom line results

Based on data from the 19 trials and 10 systematic reviews of case control/diagnostic accuracy studies included in the rapid review some key findings relating to the evidence on efficacy, harm-benefit and cost-effectiveness may be summarised as follows.

Biomarkers:

Biomarker panels tend to show better specificity in cancer detection than single markers. (Anghel et al. 2021; Carozzi et al. 2017 a/b; Chu et al. 2018; Hulstaert et al. 2021; Tarney et al. 2019).

Biomarkers not only facilitate cancer detection, but can also enhance detection of pre-cancerous lesions, e.g., Cytosponge®-TFF3 for Barrett's Oesophagus (Fitzgerald, di Pietro, O'Donovan, Maroni, et al., 2020; Fitzgerald, Di Pietro, O'Donovan, Muldrew, et al., 2020; Swart et al., 2021) and saliva cytokines for oral cancer (Chiamulera et al., 2021).

Across various cancer types, the biomarkers for colorectal cancer screening are the most intensively studied, including genomic, epigenetic and protein markers detected in blood, stool, urine and tissue (Anghel et al., 2021).

Imaging and artificial intelligence:

Novel image-enhanced endoscopy can improve early detection of upper GI-tract lesions in high-risk populations. Studies exploring detection rates as compared to standard white light imaging have suggested improvements with narrow bank imaging (Yoshida et al. 2021), blue laser imaging-bright (Dohi et al. 2019) and light linked colour imaging (Gao et al. 2021).

There is small-scale evidence for superiority of blue light imaging in bright mode over linked colour imaging in the detection rate of gastric cancer (Dohi et al. 2019) but the reverse has been demonstrated for colorectal adenomas (Ferreira et al., 2020).

Retrospective evidence, and lack of prospective evidence, suggest that current AI is not sufficiently specific to replace radiologist reading in breast screening programmes. A **systematic review** (Freeman et al. 2021) tested accuracy of standalone AI algorithms or AI-assisted radiologists to detect breast cancer in digital mammogram screening or test sets. In a retrospective evaluation including 79,910 women, 34/36 (94%) AI systems were less accurate than a single radiologist's original decision; all were less accurate than consensus of two or more radiologists. Five smaller studies (1086 women, 520 cancers) at high risk of bias and low applicability evaluated AI systems as more accurate than a single radiologist reading a test set. In three studies, AI used for triage screened out 53%, 45%, and 50% of women at low risk but also 10%, 4%, and 0% of cancers detected by radiologists.

3. Discussion

3.1 Summary

This rapid review provides evidence for the potential of new technologies in cancer screening, notably the use of biomarkers and imaging techniques. It is clear from the results of the search carried out for this review that research on imaging (including digital pathology) and biomarkers for cancer detection is a rapidly advancing field with a large number of ongoing studies (this study set is available from the authors).

The research landscape for circulating tumour DNA (ctDNA), circulating tumour cells and proteins, and DNA methylation markers for cancer screening was discussed in detail at the workshop (available on SAPEA website) where it was noted that large prospective studies are underway.

As noted in workshop 3 (available on SAPEA website) research is ongoing to test the effectiveness of AI-based cancer screening tools and explore how best to embed them into routine screening and clinical care. Two studies (outside the scope of this rapid review since not embedded in trials) were discussed that indicate that algorithms can perform as well as human radiologists. However, a recent systematic review of AI-based breast screening tools, that met the inclusion criteria, concluded that overall they were not currently sufficiently specific to replace human assessment of scans, and that more research is needed to demonstrate effectiveness, particularly in prospective real-world trials (Freeman et al., 2021). This is a promising area for future research and practice.

3.2 Strengths and limitations of this Rapid Review

3.2.1 Strengths

This review summarises a valuable sub-set of the evidence base. It emphasises the findings from studies within recent randomised and other controlled clinical trials, providing the evidence with the least potential for bias, supplemented with data from published systematic reviews of multiple case-control or diagnostic accuracy studies.

3.2.2 Limitations

In order to complete the review in a timely fashion a pragmatic and precise search strategy was employed. It is possible that additional studies within controlled trials would have been identified should there have been time for a detailed and sensitive systematic search.

It is acknowledged that other types of non-trial evidence are relevant to the topic, notably individual studies of ‘real life’ screening cohorts. In all, 101 cohort or dataset studies with ≥100 subjects were retrieved by the search and full details are available from the authors of this report.

The timeline also precluded any statistical or meta-analysis of findings. No formal critical appraisal was carried out although information is provided on whether the trial included a power calculation. Data extraction and summary were undertaken by different reviewers and, although reviewed by another author, these have not been independently checked for accuracy and consistency.

4. References:

1. Aalami, A.M., et al., *Urinary Angiogenin as a Marker for Bladder Cancer: A Meta-Analysis* Biomed Res Int, 2021. **Article ID 5557309** DOI: http://dx.doi.org/10.1155/2021/5557309.
2. Anghel, S.A., et al., *Promising epigenetic biomarkers for the early detection of colorectal cancer: A systematic review.* Cancers, 2021. **13(19) (no pagination)** DOI: http://dx.doi.org/10.3390/cancers13194965.
3. Campi, R., et al., *Novel Liquid Biomarkers and Innovative Imaging for Kidney Cancer Diagnosis: What Can Be Implemented in Our Practice Today? A Systematic Review of the Literature* Eur Urol Oncol 2021. **4**(1): p. 22-41 DOI: http://dx.doi.org/10.1016/j.euo.2020.12.011.
4. Carozzi, F., et al., *Selection of subjects at high-risk for LDCT lung cancer screening using a molecular panel: results by the ITALUNG biomarker study.* Journal of thoracic oncology, 2017. **12**(1): p. S570-S571 DOI: http://dx.doi.org/10.1016/j.jtho.2016.11.717.
5. Carozzi, F.M., et al., *Multimodal lung cancer screening using the ITALUNG biomarker panel and low dose computed tomography. Results of the ITALUNG biomarker study.* International Journal of Cancer, 2017. **141**(1): p. 94-101 DOI: http://dx.doi.org/10.1002/ijc.30727.
6. Chauvie, S., et al., *Artificial intelligence and radiomics enhance the positive predictive value of digital chest tomosynthesis for lung cancer detection within SOS clinical trial.* European Radiology, 2020. **30**(7): p. 4134-4140 DOI: http://dx.doi.org/10.1007/s00330-020-06783-z.
7. Chiamulera, M.M.A., et al., *Salivary cytokines as biomarkers of oral cancer: a systematic review and meta-analysis.* BMC Cancer, 2021. **21**(1): p. 205 DOI: http://dx.doi.org/10.1186/s12885-021-07932-3.
8. Chu, G.C.W., K. Lazare, and F. Sullivan, *Serum and blood based biomarkers for lung cancer screening: a systematic review.* BMC Cancer, 2018. **18**(1): p. 181 DOI: http://dx.doi.org/10.1186/s12885-018-4024-3.
9. Dohi, O., et al., *Blue laser imaging-bright improves the real-time detection rate of early gastric cancer: a randomized controlled study.* Gastrointestinal endoscopy, 2019. **89**(1): p. 47-57 DOI: http://dx.doi.org/10.1016/j.gie.2018.08.049.
10. Ferrándiz, L., et al., *Internet-based skin cancer screening using clinical images alone or in conjunction with dermoscopic images: a randomized teledermoscopy trial.* Journal of the American Academy of Dermatology, 2017. **76**(4): p. 676-682 DOI: http://dx.doi.org/10.1016/j.jaad.2016.10.041.
11. Ferreira, L.E., et al., *Can new image-enhancing technologies improve adenoma detection rates in cancer screening colonoscopies? A randomized clinical trial.* Gastrointestinal endoscopy, 2020. **91**(6): p. AB232-AB233 DOI: http://dx.doi.org/10.1016/j.gie.2020.03.1769.
12. Fitzgerald, R.C., et al., *Cytosponge-trefoil factor 3 versus usual care to identify Barrett's oesophagus in a primary care setting: a multicentre, pragmatic, randomised controlled trial.* Lancet, 2020. **396**(10247): p. 333-344 DOI: http://dx.doi.org/10.1016/s0140-6736(20)31099-0.
13. Fitzgerald, R.C., et al., *634 results from the Barrett's oesophagus trial 3*

(Best3): A randomised controlled trial comparing the Cytosponge™-Tff3 test with usual care to identify oesophageal precancer in primary care patients with chronic gastroesophageal reflux. Gastroenterology, 2020. **158**(6): p. S-136- DOI: http://dx.doi.org/10.1016/S0016-5085(20)31020-9.

14. Freeman, K., et al., *Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy.* BMJ, 2021. **374**: p. n1872 DOI: http://dx.doi.org/10.1136/bmj.n1872
15. Gao, J., et al., *Linked Color Imaging Can Improve Detection Rate of Early Gastric Cancer in a High-Risk Population: A Multi-Center Randomized Controlled Clinical Trial.* Digestive Diseases & Sciences, 2021. **66**(4): p. 1212-1219 DOI: http://dx.doi.org/10.1007/s10620-020-06289-0.
16. Gruner, M., et al., *Narrow-band imaging versus Lugol chromoendoscopy for esophageal squamous cell cancer screening in normal endoscopic practice: randomized controlled trial.* Endoscopy, 2021. **53**(7): p. 674-682 DOI: http://dx.doi.org/10.1055/a-1224-6822.
17. Guo, Y., J. Long, and S. Lei, *Promoter methylation as biomarkers for diagnosis of melanoma: A systematic review and meta-analysis.* Journal of Cellular Physiology, 2019. **234**(5): p. 7356-7367 DOI: http://dx.doi.org/10.1002/jcp.27495.
18. Hostetter, J.M., et al., *Personalizing lung cancer risk prediction and imaging follow-up recommendations using the National Lung Screening Trial dataset.* Journal of the American Medical Informatics Association, 2017. **24**(6): p. 1046-1051 DOI: http://dx.doi.org/10.1093/jamia/ocx012.
19. Hubers, A.J., et al., *DNA hypermethylation analysis in sputum of asymptomatic subjects at risk for lung cancer participating in the NELSON trial: argument for maximum screening interval of 2 years.* Journal of Clinical Pathology, 2017. **70**(3): p. 250-254 DOI: http://dx.doi.org/10.1136/jclinpath-2016-203734.
20. Hulstaert, E., et al., *Candidate RNA biomarkers in biofluids for early diagnosis of ovarian cancer: A systematic review.* Gynecologic Oncology, 2021. **160**(2): p. 633-642 DOI: http://dx.doi.org/10.1016/j.ygyno.2020.11.018.
21. In, H., et al., *Serum pepsinogen as a biomarker for gastric cancer: a nested case-control study using the prostate, lung, colorectal, and ovarian(PLCO) cancer screening trial data.* Journal of clinical oncology, 2021. **39**(3 SUPPL) DOI: http://dx.doi.org/10.1200/JCO.2021.39.3_suppl.188.
22. Kazarian, A., et al., *Testing breast cancer serum biomarkers for early detection and prognosis in pre-diagnosis samples.* British Journal of Cancer, 2017. **116**(4): p. 501-508 DOI: http://dx.doi.org/10.1038/bjc.2016.433.
23. Saheb Sharif-Askari, N., et al., *Integrative systematic review meta-analysis and bioinformatics identifies MicroRNA-21 and its target genes as biomarkers for colorectal adenocarcinoma.* International Journal Of Surgery, 2020. **73**: p. 113-122 DOI: https://doi.org/10.1016/j.ijsu.2019.11.017.
24. Sturgeon, S.R., et al., *White blood cell DNA methylation and risk of breast cancer in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO).* Breast Cancer Research, 2017. **19**(1): p. 94 DOI:

http://dx.doi.org/10.1186/s13058-017-0886-6.

25. Sturgeon, S.R., et al., *Prediagnostic White Blood Cell DNA Methylation and Risk of Breast Cancer in the Prostate Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) Cohort.* Cancer Epidemiology, Biomarkers & Prevention, 2021. **30**(8): p. 1575-1581 DOI: http://dx.doi.org/10.1158/1055-9965.
26. Sullivan, F.M., et al., *Earlier diagnosis of lung cancer in a randomised trial of an autoantibody blood test followed by imaging.* European Respiratory Journal, 2021. **57**(1): p. 01 DOI: http://dx.doi.org/10.1183/13993003.00670-2020
27. Sun, N., et al., *Utility of isocitrate dehydrogenase 1 as a serum protein biomarker for the early detection of non-small-cell lung cancer: A multicenter in vitro diagnostic clinical trial.* Cancer Science, 2020. **111**(5): p. 1739-1749 DOI: http://dx.doi.org/10.1111/cas.14387.
28. Swart, N., et al., *Economic evaluation of Cytosponge R-trefoil factor 3 for Barrett esophagus: A cost-utility analysis of randomised controlled trial data.* EClinicalMedicine, 2021. **37**: p. 100969 DOI: http://dx.doi.org/10.1016/j.eclinm.2021.100969.
29. Tarney, C.M., et al., *Biomarker panel for early detection of endometrial cancer in the prostate, lung, colorectal, and ovarian cancer screening trial.* Gynecologic oncology, 2019. **154**: p. 42-43 DOI: http://dx.doi.org/10.1016/j.ygyno.2019.04.102.
30. Wen, Q., et al., *Urinary Volatile Organic Compound Analysis for the Diagnosis of Cancer: A Systematic Literature Review and Quality Assessment.* Metabolites, 2020. **11**(1): p. 17 DOI: http://dx.doi.org/10.3390/metabo11010017.
31. Yang, B.H., et al., *The value of DNA quantitative cytology test for the screening of endometrial cancer.* Cancer Management and Research, 2019. **11**: p. 10383-10391 DOI: http://dx.doi.org/10.2147/CMAR.S225672.
32. Yoshida, N., et al., *Early gastric cancer detection in high-risk patients: a multicentre randomised controlled trial on the effect of second-generation narrow band imaging.* Gut, 2021. **70**(1): p. 67-75 DOI: http://dx.doi.org/10.1136/gutjnl-2019-319631.
33. Yu, Y., et al., *Long noncoding RNAs as diagnostic biomarkers for the early detection of digestive tract cancers: a systematic review and meta-analysis.* Revista Espanola de Enfermedades Digestivas, 2020. **112**(10): p. 797-804 DOI: http://dx.doi.org/10.17235/reed.2020.5450/2018.
34. Zhou, X., et al., *Association of APC gene promoter methylation and the risk of gastric cancer: A meta-analysis and bioinformatics study.* Medicine (Baltimore), 2020. **99**(16): p. e19828 DOI: http://dx.doi.org/10.1097/MD.0000000000019828.

5. Rapid review method

5.1 Eligibility criteria

Inclusion:

- Studies within a randomised controlled trial (RCT) or controlled clinical trial[2] or a systematic review published since 2017
- 'New technology' interventions including: Artificial intelligence, machine learning, genetic markers (including ctDNA, mRNA), imaging, urinary markers, f(a)ecal markers, volatile compounds, auto-antibodies.
- Screening for first (early) diagnosis of any cancer in the general population
- Inclusion of data on efficacy, harm-benefit or cost-effectiveness relating to targeted screening methods using new technology(ies)
- All locations, all languages but to emphasise the findings from EU studies within the narrative write up

Exclusion:

- Studies looking at new technologies to
    - Support decision making/informed choice
    - Aid cancer detection (post screening)
    - Aid cancer detection in symptomatic patients
    - Assess prognosis
- Studies to explore implementation factors such as adherence to testing
- HPV testing and/or further testing for those with HPV positive status (trial data in RR2)
- Helicobacter pylori testing
- MRI for breast cancer (trial data in RR2) and prostate cancer (trial data in RR1)
- Cytology for anal cancer
- Conference reports
- Non-English language studies
- Studies based on large screening cohorts or datasets

6.2 Literature search strategy

Searches were carried out for publications from 2017 onwards using title and Medical Subject Heading (MeSH) searches of the Cochrane Central Register of Controlled Trials (CCTR), Medline, Embase, the ICTRP trials register and Clinicaltrials.gov.

Search terms

[2] Quasi-randomised and other controlled trials where randomisation is not explicit, but cannot be ruled out

Text words: (cancer* AND (screen* OR early detection)) in title
MeSH terms: (exp Neoplasms/) AND (early detection of cancer/)

Combined with

(machine learning OR artificial intelligence OR biomarker* OR AI OR ctDNA OR mRNA OR microRNA OR DNA OR imaging OR urinary marker* OR faecal marker* OR fecal marker* OR VoC* OR volatile compound* OR antibod* OR anti-bod* OR cytosponge) in title

In Medline using above terms [AND randomized controlled trial.pt OR controlled clinical trial.pt OR pragmatic clinical trial.pt OR systematic review.m_titl OR exp mass screening/ OR trial.m_titl OR cohort.m_titl];

In Embase using above terms [AND exp mass screening/ OR systematic review.m+titl].
Randomized controlled trial.pt OR trial.m_titl OR cohort.m_titl

Additional search methods: The workshop on the topic was attended by one of the review authors to note any additional studies meeting the inclusion criteria.

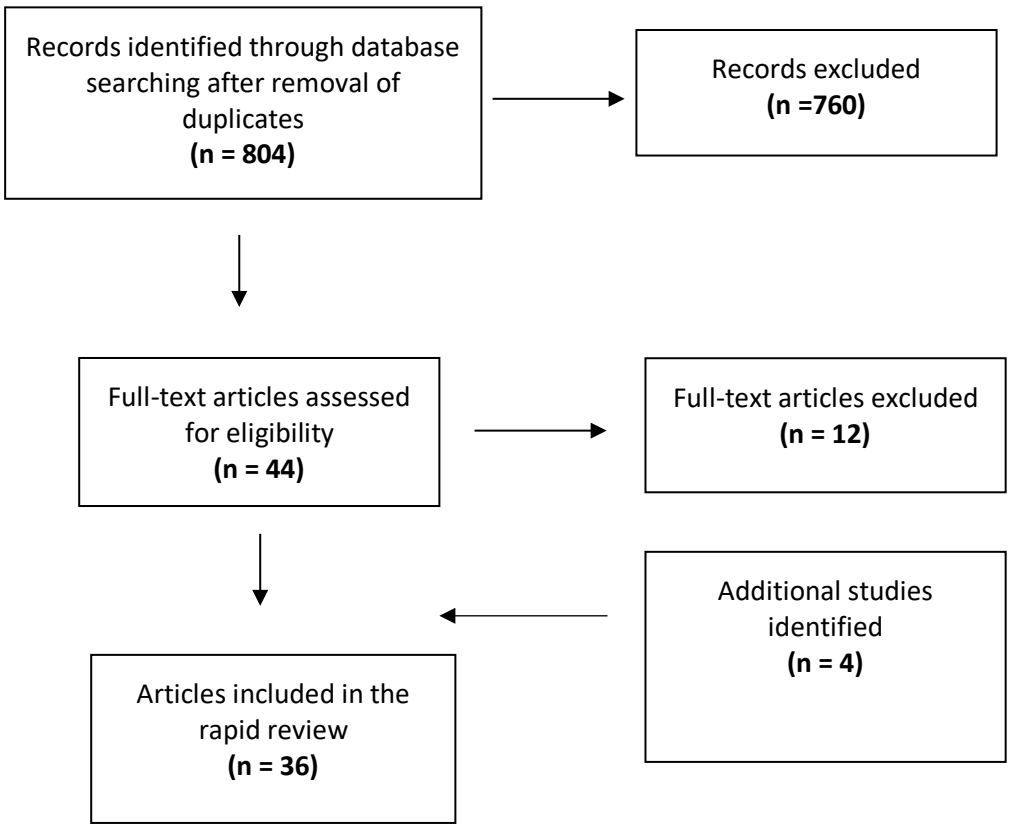
6.3 Resources list

Clinical trials.gov
Cochrane Library [Cochrane Reviews/Cochrane Central Register of Controlled trials]
Health Technology Assessment
Embase
International Clinical Trials Registry Platform (ICTRP)
Medline
US Preventive Services Taskforce (USPSTF)

6.4 Study selection process

Results from the literature searches were imported into EndNote 20, where duplicates were removed. Titles and abstracts were screened for inclusion followed by full text screening. Both screening stages were undertaken by a team of reviewers according to the eligibility criteria in Section 5.1.

5.5 Study selection flow chart



5.6 Data extraction

Data from main trial report(s) on efficacy, harm-benefit or cost effectiveness were extracted into a summary table for each cancer by a single reviewer (Section 2.1).

5.7 Quality appraisal

In this review, most of the included data was from diagnostic accuracy studies within a trial or from systematic reviews of diagnostic accuracy studies. Each included study was identified as a systematic review; or an RCT or controlled clinical trial (CCT) according to the study design as provided in the database(s) within the evidence table (Section 2.1) along with a note as to whether a power calculation was included as part of the trial. No other formal critical appraisal was carried out.

5.8 Synthesis

The findings are summarised in a narrative report, drawing from the summary tables with brief findings based on the consensus from the included studies.

6. Additional information

6.1 Conflicts of interest

None

6.2 Acknowledgements

This template is based, with permission, on the rapid review template used within the Palliative Care Evidence Review Service (PaCERS) and the Welsh Covid 19 Evidence Centre.

7. About the review team

The Specialist Unit for Review Evidence (SURE) is a team of experienced systematic reviewers and information specialists at Cardiff University who conduct all forms of systematic and other evidence reviews, and teach evidence review methods. The team work across all topic areas and also specialise in health and social care. Staff have carried out a number of reviews for SAPEA, working closely with Academia Europaea and experienced reviewers within the University's Library Service. Reviews are carried out in close collaboration with subject specialists for each review topic. For these rapid reviews the subject specialists are Dr Hui-Ling Ou (Cambridge University) and Dr Nicholas Courtier (Cardiff University).

SAPEA is part of the European Commission's Scientific Advice Mechanism, which provides independent, interdisciplinary, and evidence-based scientific advice on policy issues to the European Commission.

SAPEA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 737432.

www.sapea.info
@SAPEAnews