Early detection of cancer

Authors: David Crosby[1*], Sangeeta Bhatia[2,3], Kevin M. Brindle[4,5], Lisa M. Coussens[6,7], Caroline Dive[8,9], Mark Emberton[10], Sadik Esener[7,11,12], Rebecca C. Fitzgerald[13], Sanjiv S. Gambhir[14], Peter Kuhn[15], Timothy R. Rebbeck[16,17], Shankar Balasubramanian[18,19*]

Affiliations:

[1]Cancer Research UK, London, UK.

[2]Marble Center for Cancer Nanomedicine, Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, USA.

[3]Howard Hughes Medical Institute, Chevy Chase, USA.

[4]Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK.

[5]Department of Biochemistry, University of Cambridge, Cambridge, UK.

[6]Cell, Developmental & Cancer Biology, Oregon Health & Science University, USA.

[7]Knight Cancer Research Institute, Oregon Health & Science University, USA.

[8]Cancer Research UK Lung Cancer Centre of Excellence at Manchester and University College London, UK.

[9]CRUK Manchester Institute Cancer Biomarker Centre, University of Manchester, Manchester, UK.

[10]Division of Surgery and Interventional Science, University College London, London, UK

[11]Department of Biomedical Engineering, School of Medicine, Oregon Health & Science University, USA.

[12]Cancer Early Detection & Advanced Research Center, Oregon Health & Science University, USA.

[13]MRC Cancer Unit, Hutchison-MRC Research Centre, University of Cambridge, Cambridge, UK.

[14]Department of Radiology, Molecular Imaging Program at Stanford (MIPS), Stanford University, Stanford, USA.

[15]University of Southern California, Los Angeles, USA.

[16]Division of Population Science, Dana Farber Cancer Institute, Boston, USA.

[17]Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, USA.

[18]Department of Chemistry, University of Cambridge, Cambridge, UK.

[19]Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK.

[*]Correspondence to: david.crosby@cancer.org.uk, sb10031@cam.ac.uk

One Sentence Summary (125 characters):

A Review discussing how to overcome key challenges to enable early detection of cancer for earlier intervention and increased survival.

PRINT SUMMARY

Background

When cancer is detected at the earliest stages, treatment is more effective, and survival dramatically improves. Yet around 50% of cancers are still detected at an advanced stage. Improved earlier detection of cancer could substantially increase survival rates. Although recent advances in early detection save lives, further innovations and development of early cancer detection approaches are needed. The field is evolving rapidly due to advances in biological understanding and an increasing pace of technological progress.

Advances

We highlight five challenges facing the field, current work in those areas and where more research is needed to move early detection towards reality.

The first challenge is to build a greater understanding of the biology and behavior of early disease, which will help identify ways to distinguish between consequential, aggressive lesions and those inconsequential lesions that will not cause harm. Such insight will be crucial in realizing the potential for early detection to inform treatment decisions and impact on survival, while minimising the risk of over-treatment. Alongside studies in human samples, better models of disease are enabling identification of early signals of tumorigenesis and clarifying the contribution of the immune system and microenvironment to tumor development, which will be crucial to achieving this goal.

The second challenge is determining the risk of cancer; how can we use germline genomic susceptibility, family history, exposures and other risk factors, demographic and behavioral data to build nuanced risk models that may identify who should be tested for cancer and how test results should be interpreted and followed up. Progress is being made against this challenge through advances in understanding of the genomics of cancer risk, integration of that insight with other risk factors and the development of large-scale population cohorts where risk models can be developed and validated.

The third challenge is finding and validating biomarkers of early cancer. There is considerable difficulty in finding accurate signals of early cancer (which usually exist in very low amounts) amidst the noise of normal human physiology. While progress has historically been slow, many promising early detection markers are emerging, including circulating tumor DNA, circulating tumor cells, proteins, exosomes and cancer metabolites. Advances in data analytic methodologies (such as machine learning) and integration across marker types in multi-modal tests are also accelerating progress.

The fourth challenge is technological. It involves both the iterative improvement of existing approaches and the development of disruptive detection technologies that can very sensitively and specifically identify early biological changes, whether in tissue structure, biochemistry, or function. Powerful molecular biological analytic technologies and advanced imaging and histopathological methods are increasing the ability to sensitively find earlier tumors, while the use of synthetic markers may help to amplify their signal.

The fifth challenge is how to most appropriately evaluate early detection approaches. Translation of biological insights into new diagnostic technologies and execution of clinical trials to validate those advances requires substantial time and money. We discuss ways in which that process might be improved.

Outlook

For early detection to deliver transformative progress in cancer survival, wider skillsets beyond cancer biology are essential, e.g. engineers, chemists, physicists, technology developers, behavioral and computer scientists. Integrated, interdisciplinary collaboration will be key to bring new ideas to ultimately address the challenges of cancer early detection. We believe early detection of cancer is approaching a tipping point as biological insight and technological capacity are increasing at an unprecedented rate. This Review discusses the current state of the field and suggests constructive ways forward that build on current progress to deliver effective earlier detection of cancer and appropriate intervention.

Figure title and short legend:

Abstract (125 words):

Survival improves when cancer is detected early. However, ~50% of cancers are at an advanced stage when diagnosed. Early detection of cancer or pre-cancerous change allows early intervention to try to slow or prevent cancer development and lethality. To achieve early detection of all cancers, numerous challenges must be overcome. It is vital to better understand who is most at risk of cancer. We also need to elucidate the biology and trajectory of early- and pre-cancer to identify consequential disease that requires intervention. Insights must be translated into sensitive, specific, early detection technologies and appropriately evaluated to support practical clinical implementation. Interdisciplinary collaboration is key: advances in technology and biological understanding highlight that it is time to accelerate early detection research and transform cancer survival.

Introduction

Cancer is a major global public health problem; there were 10 million deaths from cancer worldwide in 2020 (*1*). It is the second leading cause of death globally, causing one in six deaths (*2*). For nearly all cancers, the chances of survival increase significantly if the disease is detected, diagnosed and treated at an early stage (*3*) **(Fig. 1)**.

Early detection aims to detect consequential cancer or pre-cancerous change at the earliest time-point at which intervention could improve survival or reduce morbidity. Consequential disease will cause mortality or substantial morbidity within the individual's expected remaining lifespan. Early detection can take place across several windows in the transition from normal cellular activity to dysregulation to cancer; this includes detecting cancer itself at an earlier point in its development, but also detecting precursor changes (**Fig.2**). Screening, which pro-actively tests asymptomatic people, is a subset of early detection measures. Many of the principles of early detection interact with other points in cancer care such as detection of minimal residual disease or disease recurrence (**Fig. 2**). This

Review focuses on early detection of primary cancers and pre-cancerous changes in both screening and symptomatic detection contexts.

Cancer early detection research and development has produced tremendous health benefits e.g. through established screening approaches such as for cervical, breast and colorectal cancers, which are now diagnosed less frequently at later stages than those cancers without established screening (*4*) (**Fig. 1**). But many cancers, such as esophageal, pancreatic and ovarian, are still often diagnosed at advanced stages, when prognosis is extremely poor.

While early detection confers survival advantages in all populations, approximately 70% of cancer deaths occur in low- and middle-income countries (*2*), often with late diagnosis. For example, the rate of late-stage breast cancer diagnosis in black sub-Saharan African women remained well above 60% from the 1970s to 2011, whereas in the US, that rate of late diagnosis decreased from ~60% to 32% in black women over the same period (*5*). Some cancers that have effective early detection tests, such as cervical cancer, have much higher mortality rates in low human development index (HDI) countries (19.8 per 100,000) compared with high HDI countries (3.1 per 100,000), whereas other cancers without effective early detection tests differ less (e.g., stomach cancer, 5.0 per 100,000 in high versus 4.0 in low HDI countries). Detection of cancer at late stages is a global problem that is exacerbated in resource-poor settings; equity is a considerable challenge (*6*-*8*). Patients diagnosed with later-stage cancer can miss the window for curative intervention, and expensive later-stage systemic treatments are often associated with severe side effects and worse outcomes (**Fig. 1**). Further research to build on early detection successes and extend into other cancer types could transform patient outcomes.

The challenges facing early detection research fall into five broad categories:

First, understanding the biology of early cancer; we know comparatively little about the earliest events in cancer genesis. What should we look for, and, once found, how can we

know which early lesions will progress to become aggressive, consequential disease versus indolent, inconsequential disease?

Second, determining risk; populations differ from each other, and individuals even more so. There are great challenges in knowing which populations or individuals are at greater risk of having or developing cancer, and therefore deciding who should be tested and how tests should be interpreted and acted on.

Third, finding and validating biomarkers; early tumors are miniscule – discovering sensitive markers of their presence and robustly validating them presents an archetypal needle-in-a-haystack challenge.

Fourth, developing accurate technologies; there is significant challenge in developing technologies sensitive enough to detect markers of early cancers. An equally important challenge is to ensure those technologies are also sufficiently specific to not raise false alarms or lead to over-treatment for inconsequential disease.

Fifth, evaluating early detection approaches appropriately; the ultimate challenge is to robustly demonstrate that a new early detection approach can indeed detect cancers early, and ultimately save lives. The relative scarcity of cancer in the general population can make this a difficult, prolonged and extremely expensive process.

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic has highlighted the general need for accurate, early detection technologies that address the issues of cost, access and scaling, public acceptance of testing, and integration of diagnostics with public health infrastructure and decision-making. The point-of-care tools and privacy-compliant telehealth solutions that have emerged to meet the pandemic crisis may also help advance the implementation of early cancer detection.

Early detection approaches must ensure that they address, rather than exacerbate, health inequities, and that the balance of benefit to harm (through overdiagnosis, unnecessary

invasive follow-up and overtreatment for inconsequential disease) is positive. In this Review, we describe the diverse research challenges, and propose ways toward achieving early detection of cancer.

Challenge 1: understanding the biology of early cancer

There is a continuum in tumorigenesis from normal to dysregulated to cancerous; a key challenge is to understand this biology so that we can predict the future trajectory of the changes we detect and determine when early disease becomes consequential and/or lethal.

***The cancer continuum and transition to lethality*.** Cancer evolves from early inconsequential dysregulation in molecular and cellular phenotypes, to malignant transformation where critical changes in the cell's genome or epigenome culminate in a hallmark series of abnormal features that define cancer, to potentially lethal invasion and metastasis, and ongoing cellular evolution and diversification (*9*). Windows of opportunity for cancer detection exist across this continuum, with challenges to each (Fig. 2). The transition rate through these stages depends on the cancer type and so understanding this timeline can help pinpoint the optimal time for detection and intervention.

Annual screening may not detect fast, aggressive cancers that develop between screening visits (*10*). Conversely, slow-growing cancers undergoing malignant transition over several years can be tracked with active surveillance and screening of at-risk populations. Some cancers follow a clear path from precursor condition to malignancy, such as polyps preceding colon cancer. However, not all precursors will progress to cancer, and not all cancers will be consequential. For example, the pre-cancerous condition monoclonal gammopathy of undetermined significance (MGUS) has an average risk of developing into multiple myeloma (a lethal cancer) of only 1% per year (*11*), and the risk of Barrett's

esophagus developing into cancer is 0.3% to 30% per year (*12*). We do not fully understand which lesions will progress to consequential disease, and which will not.

What confers "lethality" and its timing? A cancer can theoretically be traced, via an evolutionary tree, back to a single cell. This single cell arises from a specific set of conditions, including the tissue microenvironment and the immune system. Each organ system presents a different environment, with some mutations causing a potentially lethal tumor in one organ context but not another (*13*). The picture also changes within individuals due to ageing. A tumor-permissive environment can be created by cellular and molecular changes in non-cancerous cells during ageing, such as biophysical alterations in the extracellular matrix, changes in secreted factors and changes in the immune system (*14*). What are the transitions leading to that initial cancer cell, and then the changes that engender a consequential tumor, both within the cell and with its interactions with its microenvironment? The early evolution of most cancers cannot easily be observed in people due to clinical presentation at advanced disease stages, and tissue sampling difficulties when monitoring pre-cancers. Blood cancers are an exception, where the ease of blood sampling has allowed better understanding of clonal hematopoiesis (*15*). For example, all multiple myelomas will have progressed from MGUS. Chromosomal and other mutational changes can be monitored in MGUS patients and may highlight patients who are progressing from MGUS to smoldering myeloma to malignant multiple myeloma (*16*). Clinical trials suggest that patients who undergo early detection of MGUS progression may benefit from therapeutic intervention at the stage of smoldering myeloma, rather than waiting for symptomatic malignant myeloma with end-organ damage (*17*). This demonstrates how detection and molecular stratification of a pre-neoplastic lesion (**Fig. 2**) can trigger intervention before clinically-observed definite malignancy. Given that not all patients presenting with multiple myeloma will have a prior clinical diagnosis of MGUS, this will not catch every case,

although it does give a paradigm to study and exploit the biology of transition from pre-cancer to cancer.

As well as the biology of the affected cell itself, the transition from normal to cancer is also affected by the cell's microenvironment. The microenvironment includes host immune cells, extracellular matrix, and secreted proteins, and it is in various states of hypoxia and pH. The microenvironment surrounding the would-be tumor cell can contribute to tumor progression, determining whether a tumor cell remains localized or spreads aggressively. An early tumor may also induce detectable changes in its microenvironment, generating potential biomarkers for detection. These changes could also indicate whether an initial lesion will become consequential.

The immune system is a crucial regulator and indicator of the initiation and progression of early tumors (for example, the spatial positioning of tumor-infiltrating lymphocytes with regards to the tumor can, in some cancer types, indicate how invasive a tumor is (*18*)), and immune system markers may be used to identify residual disease post-therapy, or to predict response to therapies. However it is becoming clear that the immune system may also may be useful as an early detection paradigm unto itself. As discussed in Challenge 3 below, the very small size of the earliest tumors means that any markers they shed into the circulation will exist at very low levels, so impeding detection through test. The human immune system is an exquisitely sensitive detection apparatus and can act as a signal amplifier (each tumor cell being potentially exposed to many immune cells), which might potentially be harnessed to signal the presence of a cancer. Immune system markers currently under investigation as cancer early detection approaches include autoantibodies and T cell repertoires.

Biological models of disease. Because we cannot easily observe the first tumor cell to emerge in humans, cancer models have been developed to probe the mechanisms underlying tumor initiation (*19-22*). However, there are few models of very early cancer or premalignant disease that faithfully reproduce somatic events leading to disease in immune-competent native tissue microenvironments.

First-generation transgenic models of human cancer progression (*23*) afforded initial glimpses of tissue- and organ-specific biologies of neoplastic progression. Although such studies have revealed tumor cell-intrinsic (*24, 25*) and extrinsic characteristics (*9, 26*) that support the principal of malignancy, these models have substantial drawbacks e.g. rapid progression, and phenotypes that are frequently fully-penetrant. Therefore, these models do not accurately recapitulate human disease.

Improvements involving immune-competent mouse models of human cancer with constitutive and conditional mutations in multiple cancer-associated genes, as well as embracing tumor microenvironment and epigenetic regulators, have improved models of early cancer. For example, mouse models allowing exploration of early tumorigenesis which more closely recapitulates human disease (including immune-competent and conditional expression models) now exist for non-melanoma squamous carcinoma (*27*), pancreatic adenocarcinoma (*28*), colon cancer (*29*) and lung adenocarcinoma (*30, 31*). The next wave of model system development using approaches including circulating tumor cell patient-derived explants (*32, 33*), patient-derived xenografts and creation of complex organoids involving multiple cell types (*34*) will enable further progress. However, many patient-derived xenograft models use samples from advanced human disease implanted into immunocompromised mice, which may not reflect truly early disease processes or the important role of the immune response to early lesions. With increasing sophistication, the interplay of patient-derived models and advanced non-human model systems can provide a

path to greater understanding of early cancer biology, early detection markers, and appropriate interventions for early cancers.

Making progress against this challenge will require an integrated approach; early detection markers and tests must seek not only to detect an early cancer but to provide enough actionable information to predict the likely course of the detected lesion (consequential or inconsequential) and therefore to properly inform clinical decision making, e.g. to monitor or to treat? Investigators in this field should seek to deliver this information through investigation of not only the cancer cell itself but also of the microenvironment and immune system, and the use of sophisticated model systems alongside human samples.

Challenge 2: determining risk of cancer

In order to understand whom to test, how and when to test, and also how test results should be interpreted, requires understanding of individual cancer risk so as to maximize benefit from early detection and to minimize the risk of harm (through over- or under-diagnosis and treatment).

Early detection strategies will not be of equal value to everyone. Therefore, it is important to identify the people at elevated risk of cancer and to tailor an early detection strategy to that group to maximize the benefits and minimize the harms of early detection (*35*).

Risk models. Risk assessment models can identify individuals or populations at increased risk for a specific cancer or cancers. Risk stratification includes information about age, familial history, exposures, and lifestyle (*36*), which can be augmented by genetic screening to detect variants in genes associated with cancer. The strategy is exemplified by breast cancer risk prediction models used to stratify women into higher risk categories and towards genetic testing for inherited cancer susceptibility (*37, 38*). Women with inherited *BRCA1* or *BRCA2* pathogenic variants that are associated with increased risk of breast and ovarian cancer are candidates for chemoprevention with selective estrogen receptor modifiers, risk-reducing surgery, or enhanced breast magnetic resonance imaging (MRI) screening to enable earlier detection. Currently, very few high-risk single genes (such as *BRCA1* and *BRCA2*) trigger such action; increasingly the risk conferred by multiple genetic variants (called polygenic risk score) is being explored (*39*). The discovery and use of more informative markers of risk (be they genomic or phenotypic, e.g. breast density) integrated into models that consider family history, behavioral factors and germline variation will enable better precision in the identification of high-risk people who require screening or close surveillance and longitudinal testing for early cancer detection. It is crucial that risk models are evaluated using the appropriate methodologies (informed by statistical experts) and validated in independent data sets (*40*).

Constructing improved risk stratification models requires data and biological samples from large cohorts, ideally in pre-diagnostic populations that are followed for any cancer diagnoses. Current examples include the UK Our Future Health initiative, which will follow 5 million volunteers (https://ourfuturehealth.org.uk/), Project Baseline in the US following 10,000 volunteers (www.projectbaseline.com), the Asia Cohort Consortium following at least 1 million volunteers (www.asiacohort.org) and the European EPIC study following over 500,000 volunteers (https://epic.iarc.fr/). These longitudinal studies in healthy volunteers could help understand the hidden variability between healthy individuals and discover, validate, and contextualize early disease signals. Ultimately, these studies could identify factors to stratify healthy individuals into groups at risk of developing certain cancers.

Screening at-risk populations. Once validated risk models have identified the at-risk populations, these individuals can be invited to participate in screening programs, where available. Screening aims to detect early cancer by inviting asymptomatic, ostensibly healthy people for testing. Ideally, cancer screening should be minimally invasive or non-invasive, low cost, and provide minimal false negatives and false positives to minimize harm and maximize benefits of screening. Several existing screening tests improve cancer-specific mortality or overall mortality, including mammography for breast cancer (*41*), the Pap smear for cervical cancer (*42*), colonoscopy for colorectal cancer (*43*), and low-dose computed tomography (CT) (*44*) for lung cancer. While effective, these technologies are not necessarily minimally invasive, low cost, or highly sensitive and specific. Nor do these tests reach all of the at-risk populations concerned. For example, in the US as of 2019 < 5% of all eligible individuals have been screened for lung cancer (*45*), due to incomplete implementation of screening in health systems and low individual compliance.

For screening to be successful, the follow-up diagnostic workup must be feasible and risk-appropriate. For example, breast nodule biopsy (triggered by a positive mammogram) is a low-risk outpatient procedure. Conversely, lung biopsy (triggered by a positive lung CT screen) is highly invasive and relatively high-risk. The performance characteristics of the primary screening test, and the threshold set for a positive or negative test result, must be calibrated against the consequences of a positive result. Therefore, any early detection strategy should give rise to actionable evidence-based follow-up.

Making progress against this challenge will likely require a broader and more integrated approach than is commonly adopted; investigators should seek to develop and evaluate early detection tests within a strong context of understanding of the risk of developing cancer. This understanding needs to inform how and when tests are used, in whom, and how they are acted on. In order to best overcome this challenge, it is likely that researchers will need to increasingly adopt a multifactorial approach to understanding cancer risk, across risk factors such as the germline genome, phenotypic measures, family history, behavioral factors and more.

Challenge 3: finding and validating cancer detection biomarkers

A key challenge is how to detect the very small signal of the earliest cancers amidst the noise of normal human biology. Two fundamental measures of a diagnostic test are sensitivity and specificity. Sensitivity is the ability of a test to correctly identify those with the condition being tested for (the true positive rate); a test with higher sensitivity will miss fewer cases (false negatives). Specificity is the ability of a test to correctly identify those individuals without the condition being tested for (the true negative rate); a test with high specificity does not give a positive result when the condition is not present (false positives) (*46*). Sensitivity and specificity depend on both the technology used in the test and also the biomarker/s being measured. Two other key measures are: Positive predictive value, which is the probability that individuals that test positive actually have the disease, and negative predictive value, which is the probability that individuals who test negative do not have the disease (*47*). The target values of these parameters will depend on the intended circumstance of use of the test and also on the prevalence of the particular cancer being tested for in a given population.

Challenges in biomarker validation. Many biomarkers for early cancer detection have been proposed, but few have been validated in large trials. For example, elevated prostate specific antigen (PSA) in the blood was a candidate prostate cancer early detection biomarker. However, PSA varies greatly between individuals and within individuals as they age (or as they develop other non-malignant prostate conditions) leading to the potential for over-diagnosis, unnecessary diagnostic workup (including invasive biopsy, which confers risk), and overtreatment of inconsequential disease (which incurs potential adverse effects without increasing survival) (48, 49). As such, there has been uncertainty as to the utility of PSA as a screening tool, and it is not generally recommended as a primary, population-level screen (Figure 3). Another example of a blood marker for cancer that showed promise was CA-125 for ovarian cancer; while use of this marker increased the number of early stage diagnoses

and decreased the number of late stage diagnoses, this was not accompanied by improvement in mortality (50).

Even if validated, highly specific biomarkers can display dichotomy when taken out of context. For example, in colorectal cancer KRAS mutations are strongly associated with disease progression (*51*), but in the pancreas, many neoplasms carrying KRAS mutations are not malignant (*52*). A useful biomarker must provide enough prognostic, actionable information to inform clinical decision-making.

***Promising biomarkers*.** Biomarkers of early cancer include visible structural changes to the tissue and biochemical changes. Minimally invasive sampling methods are preferred, especially where repeated samples from healthy and at-risk people are required. In practice this includes imaging, sampling body fluids such as blood, saliva, or urine (*53*), and sampling tissues via swabs or brushings. Exhaled breath is another source of biomarkers, specifically volatile organic compound (VOC) signatures of cancer and associated metabolites (*54*).

Liquid biopsies (sampling of body fluids) can be used to identify a wide range of substances indicative of cancer, derived either from the tumor itself or from the body's response to the tumor. For example, nucleic acid fragments enter the blood during cellular apoptosis or necrosis, termed cell-free DNA (cfDNA). In cancer patients, part of the cfDNA is derived from the tumor, termed circulating tumor DNA (ctDNA). Analysis of ctDNA has shown tremendous promise for personalized mutation profiling and longitudinal monitoring of patients with advanced cancers (*53, 55*), in whom ctDNA levels are relatively high. A key challenge is that ctDNA and indeed all biochemical cancer biomarkers are present at extremely low concentrations in early stage cancer. New approaches are needed to improve on current limits of detection to address this limitation.

Human genome sequencing (*56, 57*) has provided unprecedented insights into cancer genomes (*58, 59*) and the identification of biomarkers in the genome. Although mostly focused on advanced cancer, these studies have elucidated patterns of genetic variants across

cancers, some of which may also be present in early tumors, and that can provide a basis for detection, stratification, and treatment of cancers (*60*). Liquid biopsy tests based on cancer-associated mutations in ctDNA are showing promise in early detection (*61*). However, it is increasingly clear that phenotypically normal tissue also harbors a range of somatic mutations that might normally be considered indicative of cancer or to be drivers of cancer genesis (*62*); those developing early detection approaches must be mindful of this – how can we define what a normal background of mutations is, as distinct from a consequential cancer signal?

Epigenetic modifications of DNA provide another source of early detection biomarkers. These include cancer-specific DNA methylation profiles (*63*), non-coding RNAs (*64*), small regulatory RNAs and the DNA modification 5-hydroxymethylcytosine (*65*). One promising approach analyzes methylation patterns of cfDNA in blood (*66*) and is now entering large-scale prospective clinical trials in the UK (NCT03934866) and the US (NCT04241796). Another emerging technique is based on the observation that fragmentation patterns in cfDNA differ between people with and without cancer and between different cancer types (*67*).

Other potential detection biomarkers include circulating tumor cells (*68*), exosomes (*69*), cell fusions (*70*), metabolites (*71*) and proteins (*72*). These complement DNA sequencing for the discovery and exploitation of cancer-specific signatures (*73-75*). Furthermore, certain microbes may confer susceptibility to certain cancers (*76-79*), yielding another potential pool of biomarkers.

Various types of signal modalities are in clinical use or under development for early cancer detection **(Figure 4).** It is possible that “multimodal” testing will ultimately achieve higher sensitivity and specificity for early cancer than a test that uses a single type of biomarker.

Multimodal testing can be sequential or parallel. Sequential testing cascades from tests indicating risk to confirmatory test/s of another modality. While effective (e.g.

colorectal cancer screening (**Figure 5)**), this results in long, complex diagnostic journeys. Parallel testing measures different modalities and integrates those data to provide the diagnostic signal, for example detection of the same cancer via measurement of ctDNA, metabolomics, and imaging. Parallel testing has improved the accuracy of liquid biopsy tests in blood (*75*), urine (*80*), and cervical swabs (*81*). Another approach has been to profile both ctDNA mutations and serum protein biomarkers (*75, 82*), with further improvement by also adding positron emission tomography-computed tomography (PET-CT) imaging (*83*). A prominent example of a successful multimodal cancer detection test combines an assay for fecal blood with a test for known cancer-associated DNA mutations, for improved colorectal cancer screening (*84*) over the single fecal hemoglobin test and is now in clinical use.

Data analytic methods. Novel computational tools are important to analyze, integrate, and use the data generated by diagnostics. Artificial intelligence (AI) and machine learning (ML) approaches, such as support vector machine and neural network models, can discover cancer biomarkers, detect cancer-specific signatures in high-dimensional datasets, and build prospective statistical classifiers for evaluating diagnostic performance in independent cohorts (38). Such approaches offer exciting avenues for progress but are also fraught with potential challenges, of which researchers should be mindful. Many AI and ML models are criticized for being "black box" i.e. it is not possible to explain why the features (e.g. biomarkers) have been selected by the model; the creation of full interpretable models would be advantageous. Ai and ML models are often developed (or "trained") on datasets derived from selected populations which do not represent the real population where the AI-derived test would be used, and as such, the model does not translate. Some AI and ML models are of poor design and insufficient sample size, so risking bias and overfitting. The quality of design and reporting of some trials of AI approaches can also be suboptimal, calling into question the validity of their claims; such trials may not be prospective, may be at high risk of bias, may lack appropriate transparency on data and code, may lack adequate comparator groups and may deviate from existing reporting standards. And in some cases, AI and ML methodology might simply not be advantageous over statistical methods such as logistic regression.

The way forward to address these challenges is through rigorous study design informed by appropriate statistical and methodological expertise; the right analytic tool should be selected for the intended purpose. Investigators developing early detection markers should be mindful of variability between individuals with and without cancer, the specificity of their marker and the potential for false positives/over diagnosis. A promising approach is analysis of multiple biomarkers across modalities, which are integrated to produce a pattern indicative of cancer. Robust validation of markers is key; a significant challenge for the field is securing the funding for large-scale validation studies.

Challenge 4: developing accurate technologies for early detection

Developing technologies with the sensitivity to detect the earliest tumors and the specificity to minimize false positives is a key challenge. The emergence of new technologies is enabling early cancer detection with increasing accuracy.

One early detection goal is to detect emerging solid tumors which are susceptible to therapy and unlikely to have metastasized e.g. prior to development of tumor microenvironments that support enhanced angiogenesis, and before programs suppressing anti-tumor immunity are established (*22, 26, 85*), when the tumor is roughly a millimeter in diameter (comprising 10^5 to 10^6 cells). Such early lesions. Most imaging technologies in clinical development or use cannot achieve detection at this size, but new in vivo imaging instruments such as 10.5T MRI (*86*) are continuously pushing the limits.

New technologies. Sensitivity is being improved by recent technologies that detect tumor metabolites and other secondary products (Fig. 4) that are relatively more abundant than tumor cells. This can be augmented by highly specific probes, such as tumor-specific antibodies or peptides that are radio-labelled to increase signal. Other strategies include engineered diagnostics that are selectively activated in the presence of disease, such as molecular (*80, 87*) and biological (*88*) sensors that profile the *in vivo* tumor microenvironment to generate synthetic biomarkers of disease. “Activity-based diagnostics” use enzyme activity to detect or generate exogenous biomarkers that signal the presence of cancer. For instance, nanoparticles have been developed that are cleaved by dysregulated protease activity in cancer cells or their microenvironments to generate urinary reporters (*89*), and cancer-associated enzymes can metabolize exogenous VOC probes to produce volatile reporters for non-invasive detection (*87*). New synthetic biology tools include engineered probiotic (*90*) and immune cell (*91*) diagnostics for tumor detection via amplified, activity-based readouts.

Developments in material engineering and microfabrication have yielded devices that can emulate physiological microenvironments to probe tumor biology and isolate circulating tumor cells (CTCs) and extracellular vesicles from patient samples. Notable examples include

label-free capture of CTC clusters (*68*, *92*), and ultrasensitive detection of circulating exosomes with microfluidic chips or external hardware (*93*, *94*). Miniaturization has enabled new sensing approaches using wearables and implantable devices where personalized health data can inform the prevention or interception of certain diseases (*94*). More robust integration of device engineering with downstream molecular profiling technologies will help validate the relevance of these approaches for early detection.

Imaging technologies. Contemporary imaging technologies can only visualize tumors containing over 10^9 cells; this will miss many of the smaller, earliest tumors. Imaging of tissue morphology is currently used in breast cancer screening, in the form of X-ray mammography, and low-dose CT is increasingly being used to detect early-stage lung cancer in high-risk groups (*99*). Although these techniques can be used for screening, as they are relatively quick and low cost, they are subject to limited resolution and also confer risk to the patient due to their use of ionizing radiation. More advanced imaging modalities are not currently routinely used in primary screening due to high cost and low availability. Molecular imaging technologies, such as MRI (*95*) and PET (*96*), can perform early diagnosis and staging. Enhanced variations on these technologies provide the possibility of enhanced sensitivity, specificity or PPV, for example time-of-flight PET (*97*), where transit times of the photons emitted by the object generating the image signal provide a greater signal to noise ratio, and hyperpolarized MRI (*98*), where hyperpolarised carbon-13-containing moleculaes enable the collection of perfusion and metabolic information in addition to structural imaging.

Using imaging to examine multiple properties of the lesion can enhance the detection and classification of early lesions. For example, multi-parametric MRI of the prostate provides information on prostate volume, cellularity, and vascularity; this can distinguish

benign lesions from aggressive tumors requiring intervention (*100*). Imaging has the advantage of being non-invasive and easily repeatable to detect growing tumors. For example, lung cancer screening with low-dose CT repeated over time can distinguish benign lung nodules of low malignant potential from early lung cancer nodules (*99*).

The development of computer-assisted diagnostic systems help radiologists to interpret images (*101*). Computer-driven feature extraction can exploit differences in texture and shape that the naked eye cannot see. Digital attributes of the suspect lesion are called 'radiomic' features and may contain indirect information about the underlying histopathology (*102*). This is where AI and machine learning may help detect cancer (*103*) and to predict risk of progression (*104*), although issues of transparency and reproducibility must be addressed (*105*). The application of AI in imaging will require large volumes of well-annotated image data, acquired under standardized conditions, representing all populations equitably, and made widely available via curated image repositories.

Photo-acoustic imaging exposes the region of interest to pulsed laser light of a given wavelength, generating a sound that is measured by microphones or piezoelectric sensors. The level of detail and resolution of the tissue is higher than that of all other types of imaging and is free of ionizing radiation. The challenge is depth of penetration and miniaturization for clinical use (*106*). Visible light imaging through endoscopy has been a mainstay of early detection (e.g. in the colon and lung). The emerging fluorescence endoscopy technique, along with a fluorescent molecular imaging probe, has been used for enhanced detection of lesions in patients with Barrett's esophagus (*107*) and of neoplastic polyps in the colon (*108*). Hyperpolarized MRI is another promising departure from conventional imaging. This technology allows the detection of cancers by their metabolic rather than morphological differences (*109*).

Histopathology and AI. Following initial detection by biomarkers and/or imaging, histopathology is a key confirmatory diagnostic and prognostic stage of the early detection paradigm. The application of machine learning techniques to digitized slides can increase sensitivity, reduce subjectivity and inter-reader variation and predict prognosis, recurrence and tumor susceptibility to treatment (*110*). In some cases, such as Barrett's esophagus dysplasia, bowel polyps and cervical neoplasia, pathologists examine a pre-cancerous condition with the aim of identifying the transition to early cancer. Digital pathology and AI could help improve test turnaround times and diagnostic accuracy, detecting early signs of cancer and providing data for further research (*111*). Current challenges in digital pathology include handling artefacts, overcoming sample variability, lack of binary variables where a diagnosis may require a risk score and combining samples across multiple sites and cohorts.

The fundamental challenge is to reduce the limits of detection such that the earliest tumors can be identified while minimizing false positive test results. Ways forward to address this challenge through technology include approaches to amplify biological signals, or to enable continuous monitoring, for example through wearable or implantable sensors. Advanced imaging and other noninvasive approaches also may lead towards a biopsy-free model of detection and characterization, so increasing public acceptability.

Challenge 5: evaluating early detection approaches

There are many challenges around the design and methodology of trials of early detection approaches. Trials must be carefully designed to address the relevant population and measure the appropriate endpoints in order to provide statistically robust evidence to change practice. Early detection trials differ from the better-known clinical trials for therapeutics and require specialist statistical expertise to inform study design and appropriately powered sample size. For example, early detection trials' statistical power is affected by factors which do not exist in therapeutic trials, such as the number of times an individual is tested, the time between tests and the ages at which testing will be applied. However, the main challenges to the delivery of early detection trials are in their scale and interpretation.

The scale of early detection trials. Currently, regulatory or reimbursement decisions on the adoption of cancer screening tests are generally based on impact on mortality; does the use of the screening test mean fewer deaths from cancer than in an unscreened population? Demonstrating this requires very large numbers of participants (given the comparatively low incidence rate of cancers in an asymptomatic population) and very long timelines (given the potential lag between commencement of the trial, a given individual developing cancer and that cancer resulting in death). For example, the trials assessing low-dose CT screening for lung cancer in heavy smokers took 7 years and 53,454 participants in the US (*112*) and over 10 years with 15,789 participants in Europe (*113*). In a more general population (lacking the greatly increased cancer risk of heavy smoking), even greater numbers of participants are needed. For example, trials assessing screening for ovarian and prostate cancers involved over 200,000 women (*114*) and 184,000 men, respectively (*115*). This scale makes most early detection trials multicenter by default. Clinical trial networks such as those sponsored by the European Organization for Research and Treatment of Cancer (EORTC) and the US-based National Cancer Institute (NCI) can help improve research capacity in trial sites, to facilitate and accelerate such large trials.

Another attractive option is embedding research into screening programmes, which can take advantage of existing screening infrastructure. This can, for example, be done using the stepped-wedge design, where initially observations are collected during a baseline period in which no participants are exposed to the intervention (i.e. the screening test under investigation). Following this, at regular intervals (or "steps") participants (or groups of participants) are randomized to receive the intervention; these ascending steps continue until all participants have received the intervention (*116*).

One way to decrease the length and size of trials is to power the study to detect changes in surrogate endpoints (e.g. a reduction in the absolute number of late-stage diagnoses versus controls) rather than mortality (*117*). Such trials are faster and require fewer participants to record enough events in a limited timeframe. However, most healthcare systems, regulatory agencies, and guideline bodies still require evidence of reduced mortality before approving tests for marketing, reimbursement or widespread use. Advice should be sought from the relevant agencies on which surrogate endpoints might be acceptable. Studies should be designed and powered, and endpoints chosen, based on the objectives of the study

(e.g. initial signal-finding trial versus technology validation versus confirmatory prospective trial) and the intended circumstances of use. Early detection technology must generate the evidence that is required by regulators, advisory bodies, and payers for research to achieve clinical impact; proper validation (*118*) and consideration of the pathway to implementation are crucial.

Interpreting trial results. Clinical trial results of early detection technologies should be interpreted taking into consideration spectrum and lead-time biases. Spectrum bias arises when tests are assessed in a population that does not reflect the intended target population (*119*). For example, comparing a study population with established and advanced disease to a healthy control population (often young and without other chronic diseases which increase variability in the general population), which can confound the specificity of the test. Spectrum bias also arises when tests are developed using an at-risk population (e.g. heavy smokers) with high disease incidence, but the test is intended for use in the general population (with lower incidence). Such a test will lose sensitivity and even specificity in the real-world target population, which has lower prevalence of disease and other confounders. This can cause false positives and even over-diagnosis.

Lead-time bias describes the time from early detection of disease to clinical presentation of signs and symptoms (when diagnosis would otherwise have taken place) (*120, 121*). This makes survival seem longer when you detect cancer earlier by "artificially" moving the starting block back in time, even if early detection did not affect the point at which the individual died.

Spectrum bias can be addressed by validating markers and tests in populations which appropriately represent the population of intended use of the test. Lead time bias is a more complex issue; currently the method to address this bias is to conduct a trial designed to assess impact on mortality (e.g. are there fewer deaths overall in the screened group than in the unscreened group), however this then leads to the challenge of huge sample size and cost, as discussed above. The way forward to address this challenge is through careful study design; dedicated experts in screening/diagnostic methodology must be involved and the

intended target audience for the results (e.g. regulatory and guideline-developing bodies) must be consulted, when designing trials to evaluate early detection approaches. If we are to increase the glacial pace at which new early detection/screening approaches are evaluated and reach the clinic, a re-think of the evidence threshold for adoption is required, perhaps based on an absolute reduction in late-stage diagnoses (or other well-validated surrogate outcomes), with mortality data then gathered post-implementation.

Conclusions

Early detection of cancer has the potential to transform patient survival and is increasingly recognized as an area of unmet need by the public, patients, policy makers and research funders. Many of these challenges will require a sustained effort to find practical, long-term solutions – we have suggested a framework that we believe will meaningfully accelerate progress (**Figure 6**). Several contextual issues must also be carefully considered to maximize the translation of early detection research into clinical impact.

Some funders of academic cancer research have invested in programs to specifically address early detection (e.g. the US National Cancer Institute (*122, 123*) or Cancer Research UK (*124, 125*)), but the proportion of overall cancer research funding attributable to early detection remains disproportionately low, in light of the potential health benefits. More must be done, particularly in supporting validation of markers and tests (*118*). A key part of this is the need to attract early career researchers to the field and enabling them to become established. The relatively long timelines of early detection research and test development necessitate a re-think about traditional fellowship/grant models of supporting and evaluating early career researchers to incentivize them to establish a career in this field.

Furthermore, the pharmaceutical industry has proportionally invested little in early detection, as compared to the billions spent on drug development, often due to a historical

perception of an unattractive business model. However, there now appears to be an inflection (*126*) whereby investors and large corporations are increasingly willing to invest in this space (*127, 128*) and it is possible that there is a growing realization that early detection will change the business model for cancer treatment.

An interdisciplinary culture is essential to early detection research and development, which inherently needs a convergence of biological understanding, clinical insight, technology innovation, data science, risk stratification and health systems research. Without any of these essential components, the goal of transforming cancer survival cannot be realized. The implementation of interdisciplinarity can be fostered by research funders.

To have a meaningful impact on survival, early detection must be integrated into healthcare systems and must lead to evidence-based early interventions, either to prevent progression or to cure cancer. Lastly, and crucially, researchers must keep in mind that early detection should be accessible to all based on need, must not exacerbate health inequities, and must seek to do no harm (minimizing overdiagnosis and over-treatment).

With the ever-increasing depth of biological insight and an ever-increasing pace of technological innovation, we are at the tipping point for early cancer detection research and its translation to the ultimate objective of early curative interventions and increased cancer survival.

Acknowledgements

We dedicate this paper to our friend, colleague and mentor Sam Gambhir. He showed us the path to early detection of cancer and taught us to always look to the horizon. We continue to walk that path in his memory.

Paid editorial assistance was provided by Dr Fiona Dunlevy, Axcience Medical Writing. The authors also thank Dr Alexis Webb, Cancer Research UK for her editorial and logistical assistance during the preparation of this manuscript. We also acknowledge the Cancer Research UK Policy and Information Directorate for their assistance on UK cancer diagnosis and screening statistics. We also acknowledge the Early Detection, Prevention and Infection branch of the International Agency for Research on Cancer for their assistance on Asian cancer survival statistics.

Funding: Not applicable

ME receives research support from the United Kingdom's National Institute of Health Research (NIHR) UCLH/UCL Biomedical Research Centre. He was awarded NIHR Senior Investigator status in 2013.

Author contributions: a complete list of contributions to the paper [following the CRediT model]

Conceptualization – SBa, SBh, DC, CD, ME, SE, RCF, SSG

Writing – SBa, SBh, KMB, LMC, DC, ME, SE, RCF, PK, TRR

Competing interests: competing interests (including but not limited to patents, financial holdings, professional affiliations, advisory positions, and board memberships) of any of the authors must be listed (all authors must also fill out a separate, internal Conflict of Interest form). Where authors have no competing interests, this should also be declared (e.g., "Authors declare no competing interests.").

SBa is a founder and shareholder of Cambridge Epigenetix Ltd.

SBh is Board Member of Brown University and Vertex Pharmaceuticals*, holds advisory roles at Chan Zuckerberg Biohub, Global Oncology, Inc., Cristal Therapeutics, Glympse Bio*, CEND Therapeutics*, Satellite Bio*, consults for Moderna Therapeutics, and performs sponsored research for Johnson & Johnson. *SBh holds equity in these institutions.

KB holds patents with GE Healthcare on some aspects of hyperpolarization technology and is currently a consultant for SOTIO a.s. until July 31st, 2020. KMB is a member of the scientific advisory board of NVision Imaging Technologies.

LMC is a paid consultant for Cell Signaling Technologies, receives reagent and/or research support from Plexxikon Inc., Pharmacyclics, Inc., Acerta Pharma, LLC, Deciphera Pharmaceuticals, LLC, Genentech, Inc., Roche Glycart AG, Syndax Pharmaceuticals Inc., Innate Pharma, and NanoString Technologies, and is a member of the Scientific Advisory Boards of Syndax Pharmaceuticals, Carisma Therapeutics, Zymeworks, Inc, Verseau Therapeutics, Cytomix Therapeutics, Inc., and Kineta Inc.

ME is investigator / advisor / lecturer to the following companies: STEBA Biotech; Sonacare Inc; Angiodynamics Inc.

RCF is named on patents related to Cytosponge and associated assays which has been licensed by the Medical Research Council to Covidien GI Solutions (now Medtronic). RCF is a shareholder and consultant for an Early Detection company Cyted Ltd.

SSG was a Founder, Board Member and Stockholder of EARLI, Inc.

PK is a named inventor on patents relating to liquid biopsy technology, which have been licensed to Epic Sciences by The Scripps Research Institute for commercial development. PK is a named inventor on patents relating to human performance assessment, which have been licensed to Cansera for commercial development. PK is a shareholder of and consultant to Epic Sciences and Cansera. PK receives research funding from Kite, Gilead, Amgen and Novartis. PK is a member of the CRUK Early Detection committee.

FIGURES

Fig. 1. Patients survive longer when cancer is detected at an early stage. 5-year survival data for bowel, breast, lung, ovarian, and esophageal cancer, and melanoma by stage of diagnosis from Public Health England (*129*) (A) and US Surveillance, Epidemiology and End Results (SEER) database (https://seer.cancer.gov) (B). Data from the International Agency for Research on Cancer (https://survcan.iarc.fr/indexsurvcan1.php) shows 5-year survival by stage of diagnosis for colon and breast cancers in Asian countries (C). International comparison (International Cancer Benchmarking Partnership data; https://gco.iarc.fr/survival/survmark) (D) across countries for 5-year survival of colon cancer shows similar trends in percentages of patients surviving early-stage compared to late-stage disease.

Fig. 2. Windows for early detection across the course of cancer progression. Cancer evolves through various stages, offering multiple windows for early detection. Detection at each stage presents different information and choices, with the consequences of detection dependent on the level of information provided by the subsequent test(s) and the level of certainty around whether the disease will be consequential.

Figure 3. Prostate cancer detection is a cautionary tale for over diagnosis and over treatment

There are various consequences of prostate-specific antigen (PSA) testing when it is used as a screening tool. A raised PSA level is not considered useful for prostate cancer screening due to false positives and detection of inconsequential cancers that will not cause harm in the individual's lifetime. Subsequent biopsy is also imperfect because it does not always capture the tumor and may not distinguish indolent from aggressive cancers. The introduction of imaging biomarkers (e.g. multi-parametric magnetic resonance imaging - mpMRI) combined with pathology (at the junctures indicated by the + symbol in the figure) has improved prognosis through better stratification of disease.

Fig. 4. Modalities of early cancer detection. There are a wide variety of biomarker modalities that are in use, or are being developed, for the early detection and diagnosis of cancer, alongside biomarkers that can be used to guide prognosis and monitoring of treatment response or recurrence. This figure presents some of the main examples. CT: Computed

Tomography. MRI: Magnetic Resonance Imaging. FOBT: Fecal Occult Blood Test. FIT: Fecal Immunochemical Test. PET: Positron Emission Tomography. PSA: Prostate Specific Antigen. ctDNA: Circulating tumor DNA. CTC: Circulating Tumor Cell.

Figure 5. Colorectal cancer screening is an early detection success story.
Screening for colorectal cancer (CRC) relies on a cascade of diagnostic tests (A – fecal screening to endoscopy to biopsy/histopathology)) that can lead to the detection of cancers at an earlier stage (B). This has transformed CRC into a treatable cancer with increased survival rates when the cancer is caught early (see Figure 1). Population screening programs have relied on fecal occult blood tests (FOBT) that measure gastrointestinal bleeding. In the UK, FOBT was recently replaced by the fecal immunohistochemical test (FIT – a more accurate method of detecting blood in feces) and in the US, screening also includes the Cologuard FIT-DNA test (which looks for cancer-associated DNA mutations in the feces, in addition to the FIT component). Positive results from fecal screening tests usually then cascade to endoscopic examination and where appropriate, intervention. The majority of CRC, however, is still diagnosed through presentation to primary care and urgent referral routes, where symptomatic presentation is often associated with later disease stage.

Fig. 6. Overcoming barriers to enable early detection. There are system-wide challenges (grey brick) that must be tackled to reach the goal of earlier cancer detection. The multiple facets of these challenges require a diverse set of approaches and enablers (light grey) and communities (colored outer segments) to overcome them.

References

1. H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, (2021).
2. World Health Organisation, Cancer. (2021) Available at: https://www.who.int/news-room/fact-sheets/detail/cancer Accessed: 23 March 2021
3. Office for National statistics UK, "Cancer survival by stage at diagnosis for England (experimental statistics): Adults diagnosed 2012, 2013 and 2014 and followed up to 2015," (2016) https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancersurvivalbystageatdiagnosisforenglandexperimentalstatistics/adultsdiagnosed20122013and2014andfollowedupto2015.
4. R. L. Siegel, K. D. Miller, S. A. Fedewa, D. J. Ahnen, R. G. S. Meester, A. Barzi, A. Jemal, Colorectal cancer statistics, 2017. *CA: a cancer journal for clinicians* **67**, 177-193 (2017).
5. E. Jedy-Agba, V. McCormack, C. Adebamowo, I. Dos-Santos-Silva, Stage at diagnosis of breast cancer in sub-Saharan Africa: a systematic review and meta-analysis. *The Lancet. Global health* **4**, e923-e935 (2016).

6. E. Jedy-Agba, V. McCormack, O. Olaomi, W. Badejo, M. Yilkudi, T. Yawe, E. Ezeome, I. Salu, E. Miner, I. Anosike, S. N. Adebamowo, B. Achusi, I. Dos-Santos-Silva, C. Adebamowo, Determinants of stage at diagnosis of breast cancer in Nigerian women: sociodemographic, breast cancer awareness, health care access and clinical factors. *Cancer causes & control : CCC* **28**, 685-697 (2017).
7. A. M. Nelson, D. A. Milner, T. R. Rebbeck, Y. Iliyasu, Oncologic Care and Pathology Resources in Africa: Survey and Recommendations. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **34**, 20-26 (2016).
8. I. O. Morhason-Bello, F. Odedina, T. R. Rebbeck, J. Harford, J. M. Dangou, L. Denny, I. F. Adewole, Challenges and opportunities in cancer control in Africa: a perspective from the African Organisation for Research and Training in Cancer. *The Lancet. Oncology* **14**, e142-151 (2013).
9. D. Hanahan, R. A. Weinberg, Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
10. N. Pashayan, P. D. P. Pharoah, The challenge of early detection in cancer. *Science (New York, N.Y.)* **368**, 589-590 (2020).
11. R. A. Kyle, T. M. Therneau, S. V. Rajkumar, D. R. Larson, M. F. Plevak, J. R. Offord, A. Dispenzieri, J. A. Katzmann, L. J. Melton, 3rd, Prevalence of monoclonal gammopathy of undetermined significance. *The New England journal of medicine* **354**, 1362-1369 (2006).
12. Y. Peters, A. Al-Kaabi, N. J. Shaheen, A. Chak, A. Blum, R. F. Souza, M. Di Pietro, P. G. Iyer, O. Pech, R. C. Fitzgerald, P. D. Siersema, Barrett oesophagus. *Nature reviews. Disease primers* **5**, 35 (2019).
13. K. M. Haigis, K. Cichowski, S. J. Elledge, Tissue-specificity in cancer: The rule, not the exception. *Science (New York, N.Y.)* **363**, 1150 (2019).
14. M. Fane, A. T. Weeraratna, How the ageing microenvironment influences tumour progression. *Nature reviews. Cancer* **20**, 89-106 (2020).
15. C. J. Watson, A. L. Papula, G. Y. P. Poon, W. H. Wong, A. L. Young, T. E. Druley, D. S. Fisher, J. R. Blundell, The evolutionary dynamics and fitness landscape of clonal hematopoiesis. *Science (New York, N.Y.)* **367**, 1449-1454 (2020).
16. S. Manier, K. Z. Salem, J. Park, D. A. Landau, G. Getz, I. M. Ghobrial, Genomic complexity of multiple myeloma and its clinical implications. *Nature reviews. Clinical oncology* **14**, 100-113 (2017).
17. M. V. Mateos, J. F. San Miguel, Treatment for high-risk smoldering myeloma. *The New England journal of medicine* **369**, 1764-1765 (2013).
18. T. J. Honkanen, T. Moilanen, P. Karihtala, S. Tiainen, P. Auvinen, J. P. Vayrynen, M. Makinen, J. P. Koivunen, Prognostic and predictive role of spatially positioned tumour infiltrating lymphocytes in metastatic HER2 positive breast cancer treated with trastuzumab. *Scientific reports* **7**, 18027 (2017).
19. M. DuPage, T. Jacks, Genetically engineered mouse models of cancer reveal new insights about the antitumor immune response. *Current opinion in immunology* **25**, 192-199 (2013).
20. B. Olson, Y. Li, Y. Lin, E. T. Liu, A. Patnaik, Mouse Models for Cancer Immunotherapy Research. *Cancer discovery* **8**, 1358-1365 (2018).
21. D. Bedognetti, M. Ceccarelli, L. Galluzzi, R. Lu, K. Palucka, J. Samayoa, S. Spranger, S. Warren, K. K. Wong, E. Ziv, D. Chowell, L. M. Coussens, D. D. De Carvalho, D. G. DeNardo, J. Galon, H. L. Kaufman, T. Kirchhoff, M. T. Lotze, J. J. Luke, A. J. Minn, K. Politi, L. D. Shultz, R. Simon, V. Thórsson, J. B. Weidhaas, M. L. Ascierto, P. A. Ascierto, J. M. Barnes, V. Barsan, P. K. Bommareddy, A. Bot, S. E. Church, G. Ciliberto, A. De Maria, D. Draganov, W. S. Ho, H. M. McGee, A. Monette, J. F. Murphy, P. Nisticò, W. Park, M. Patel, M. Quigley, L. Radvanyi, H. Raftopoulos, N. P. Rudqvist, A. Snyder, R. F. Sweis, S. Valpione, R. Zappasodi, L. H. Butterfield, M. L. Disis, B. A. Fox, A. Cesano, F. M. Marincola, Toward a comprehensive view of cancer immune responsiveness: a synopsis from the SITC workshop. *J Immunother Cancer* **7**, 131 (2019).
22. A. K. Palucka, L. M. Coussens, The Basis of Oncoimmunology. *Cell* **164**, 1233-1247 (2016).
23. D. Hanahan, E. F. Wagner, R. D. Palmiter, The origins of oncomice: a history of the first transgenic mice genetically engineered to develop cancer. *Genes Dev* **21**, 2258-2270 (2007).

24. D. Hanahan, R. A. Weinberg, The hallmarks of cancer. *Cell* **100**, 57-70 (2000).
25. F. McCormick, Signalling networks that cause cancer. *Trends Cell Biol* **9**, M53-56 (1999).
26. D. Hanahan, L. M. Coussens, Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* **21**, 309-322 (2012).
27. N. Oshimori, D. Oristian, E. Fuchs, TGF-β promotes heterogeneity and drug resistance in squamous cell carcinoma. *Cell* **160**, 963-976 (2015).
28. S. R. Hingorani, E. F. Petricoin, A. Maitra, V. Rajapakse, C. King, M. A. Jacobetz, S. Ross, T. P. Conrads, T. D. Veenstra, B. A. Hitt, Y. Kawaguchi, D. Johann, L. A. Liotta, H. C. Crawford, M. E. Putt, T. Jacks, C. V. Wright, R. H. Hruban, A. M. Lowy, D. A. Tuveson, Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer cell* **4**, 437-450 (2003).
29. K. M. Haigis, K. R. Kendall, Y. Wang, A. Cheung, M. C. Haigis, J. N. Glickman, M. Niwa-Kawakita, A. Sweet-Cordero, J. Sebolt-Leopold, K. M. Shannon, J. Settleman, M. Giovannini, T. Jacks, Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. *Nature genetics* **40**, 600-608 (2008).
30. E. L. Jackson, N. Willis, K. Mercer, R. T. Bronson, D. Crowley, R. Montoya, T. Jacks, D. A. Tuveson, Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes & development* **15**, 3243-3248 (2001).
31. A. L. Dooley, M. M. Winslow, D. Y. Chiang, S. Banerji, N. Stransky, T. L. Dayton, E. L. Snyder, S. Senna, C. A. Whittaker, R. T. Bronson, D. Crowley, J. Barretina, L. Garraway, M. Meyerson, T. Jacks, Nuclear factor I/B is an oncogene in small cell lung cancer. *Genes & development* **25**, 1470-1475 (2011).
32. S. C. Williamson, R. L. Metcalf, F. Trapani, S. Mohan, J. Antonello, B. Abbott, H. S. Leong, C. P. Chester, N. Simms, R. Polanski, D. Nonaka, L. Priest, A. Fusi, F. Carlsson, A. Carlsson, M. J. Hendrix, R. E. Seftor, E. A. Seftor, D. G. Rothwell, A. Hughes, J. Hicks, C. Miller, P. Kuhn, G. Brady, K. L. Simpson, F. H. Blackhall, C. Dive, Vasculogenic mimicry in small cell lung cancer. *Nature communications* **7**, 13322 (2016).
33. A. Lallo, M. W. Schenk, K. K. Frese, F. Blackhall, C. Dive, Circulating tumor cells and CDX models as a tool for preclinical drug development. *Translational lung cancer research* **6**, 397-408 (2017).
34. M. Breitenbach, J. Hoffmann, Editorial: Cancer Models. *Front Oncol* **8**, 401 (2018).
35. T. R. Rebbeck, K. Burns-White, A. T. Chan, K. Emmons, M. Freedman, D. J. Hunter, P. Kraft, F. Laden, L. Mucci, G. Parmigiani, D. Schrag, S. Syngal, R. M. Tamimi, K. Viswanath, M. B. Yurgelun, J. E. Garber, Precision Prevention and Early Detection of Cancer: Fundamental Principles. *Cancer discovery* **8**, 803-811 (2018).
36. M. Jansen, M. Banks, Early detection and risk stratification of gastric cancer are likely to be refined with biopsies targeted through high-resolution-enhanced imaging. *Gut*, (2019).
37. A. Lee, N. Mavaddat, A. N. Wilcox, A. P. Cunningham, T. Carver, S. Hartley, C. Babb de Villiers, A. Izquierdo, J. Simard, M. K. Schmidt, F. M. Walter, N. Chatterjee, M. Garcia-Closas, M. Tischkowitz, P. Pharoah, D. F. Easton, A. C. Antoniou, BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genetics in medicine : official journal of the American College of Medical Genetics* **21**, 1708-1718 (2019).
38. J. Tyrer, S. W. Duffy, J. Cuzick, A breast cancer prediction model incorporating familial and personal risk factors. *Statistics in medicine* **23**, 1111-1130 (2004).
39. S. A. Lambert, G. Abraham, M. Inouye, Towards clinical utility of polygenic risk scores. *Human molecular genetics* **28**, R133-r142 (2019).
40. E. W. Steyerberg, A. J. Vickers, N. R. Cook, T. Gerds, M. Gonen, N. Obuchowski, M. J. Pencina, M. W. Kattan, Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology (Cambridge, Mass.)* **21**, 128-138 (2010).
41. A. Bleyer, H. G. Welch, Effect of three decades of screening mammography on breast-cancer incidence. *The New England journal of medicine* **367**, 1998-2005 (2012).
42. A. C. Perkins, E. N. Skinner, A Review of the Current Cervical Cancer Screening Guidelines. *North Carolina medical journal* **77**, 420-422 (2016).

43. R. Nishihara, K. Wu, P. Lochhead, T. Morikawa, X. Liao, Z. R. Qian, K. Inamura, S. A. Kim, A. Kuchiba, M. Yamauchi, Y. Imamura, W. C. Willett, B. A. Rosner, C. S. Fuchs, E. Giovannucci, S. Ogino, A. T. Chan, Long-term colorectal-cancer incidence and mortality after lower endoscopy. *The New England journal of medicine* **369**, 1095-1105 (2013).
44. K. L. Huang, S. Y. Wang, W. C. Lu, Y. H. Chang, J. Su, Y. T. Lu, Effects of low-dose computed tomography on lung cancer screening: a systematic review, meta-analysis, and trial sequential analysis. *BMC pulmonary medicine* **19**, 126 (2019).
45. M. Triplette, J. H. Thayer, S. N. Pipavath, K. Crothers, Poor Uptake of Lung Cancer Screening: Opportunities for Improvement. *Journal of the American College of Radiology : JACR* **16**, 446-450 (2019).
46. J. Shreffler, M. R. Huecker, in *StatPearls*. (StatPearls Publishing

47. A. M. Molinaro, Diagnostic tests: how to estimate the positive predictive value. *Neuro-oncology practice* **2**, 162-166 (2015).
48. D. Ilic, M. M. Neuberger, M. Djulbegovic, P. Dahm, Screening for prostate cancer. *The Cochrane database of systematic reviews*, Cd004720 (2013).
49. D. C. Grossman, S. J. Curry, D. K. Owens, K. Bibbins-Domingo, A. B. Caughey, K. W. Davidson, C. A. Doubeni, M. Ebell, J. W. Epling, Jr., A. R. Kemper, A. H. Krist, M. Kubik, C. S. Landefeld, C. M. Mangione, M. Silverstein, M. A. Simon, A. L. Siu, C. W. Tseng, Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement. *Jama* **319**, 1901-1913 (2018).
50. Menon U, Gentry-Maharaj A, Burnell M, Singh N, Ryan A, Karpinskyj C, Carlino G, Taylor J, Massingham SK, Raikou M, Kalsi JK, Woolas R, Manchanda R, Arora R, Casey L, Dawnay A, Dobbs S, Leeson S, Mould T, Seif MW, Sharma A, Williamson K, Liu Y, Fallowfield L, McGuire AJ, Campbell S, Skates SJ, Jacobs IJ, Parmar M. Ovarian cancer population screening and mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet* **397**, 2182-2193 (2021).
51. A. T. Boutin, W. T. Liao, M. Wang, S. S. Hwang, T. V. Karpinets, H. Cheung, G. C. Chu, S. Jiang, J. Hu, K. Chang, E. Vilar, X. Song, J. Zhang, S. Kopetz, A. Futreal, Y. A. Wang, L. N. Kwong, R. A. DePinho, Oncogenic Kras drives invasion and maintains metastases in colorectal cancer. *Genes & development* **31**, 370-382 (2017).
52. M. P. di Magliano, C. D. Logsdon, Roles for KRAS in pancreatic tumor development and progression. *Gastroenterology* **144**, 1220-1229 (2013).
53. E. Heitzer, I. S. Haque, C. E. S. Roberts, M. R. Speicher, Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nature reviews. Genetics* **20**, 71-88 (2019).
54. S. X. Antoniou, E. Gaude, M. Ruparel, M. P. van der Schee, S. M. Janes, R. C. Rintoul, C. I. D. G. The Lu, The potential of breath analysis to improve outcome for patients with lung cancer. *Journal of breath research* **13**, 034002 (2019).
55. V. A. Adalsteinsson, G. Ha, S. S. Freeman, A. D. Choudhury, D. G. Stover, H. A. Parsons, G. Gydush, S. C. Reed, D. Rotem, J. Rhoades, D. Loginov, D. Livitz, D. Rosebrock, I. Leshchiner, J. Kim, C. Stewart, M. Rosenberg, J. M. Francis, C. Z. Zhang, O. Cohen, C. Oh, H. Ding, P. Polak, M. Lloyd, S. Mahmud, K. Helvie, M. S. Merrill, R. A. Santiago, E. P. O'Connor, S. H. Jeong, R. Leeson, R. M. Barry, J. F. Kramkowski, Z. Zhang, L. Polacek, J. G. Lohr, M. Schleicher, E. Lipscomb, A. Saltzman, N. M. Oliver, L. Marini, A. G. Waks, L. C. Harshman, S. M. Tolaney, E. M. Van Allen, E. P. Winer, N. U. Lin, M. Nakabayashi, M. E. Taplin, C. M. Johannessen, L. A. Garraway, T. R. Golub, J. S. Boehm, N. Wagle, G. Getz, J. C. Love, M. Meyerson, Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. *Nature communications* **8**, 1324 (2017).
56. C. International Human Genome Sequencing, Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931-945 (2004).
57. D. R. Bentley, S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, H. R. Bignell, J. M. Boutell, J. Bryant, R. J. Carter, R. Keira Cheetham, A. J. Cox, D. J. Ellis, M. R. Flatbush, N. A. Gormley, S. J. Humphray, L. J.

Irving, M. S. Karbelashvili, S. M. Kirk, H. Li, X. Liu, K. S. Maisinger, L. J. Murray, B. Obradovic, T. Ost, M. L. Parkinson, M. R. Pratt, I. M. Rasolonjatovo, M. T. Reed, R. Rigatti, C. Rodighiero, M. T. Ross, A. Sabot, S. V. Sankar, A. Scally, G. P. Schroth, M. E. Smith, V. P. Smith, A. Spiridou, P. E. Torrance, S. S. Tzonev, E. H. Vermaas, K. Walter, X. Wu, L. Zhang, M. D. Alam, C. Anastasi, I. C. Aniebo, D. M. Bailey, I. R. Bancarz, S. Banerjee, S. G. Barbour, P. A. Baybayan, V. A. Benoit, K. F. Benson, C. Bevis, P. J. Black, A. Boodhun, J. S. Brennan, J. A. Bridgham, R. C. Brown, A. A. Brown, D. H. Buermann, A. A. Bundu, J. C. Burrows, N. P. Carter, N. Castillo, E. C. M. Chiara, S. Chang, R. Neil Cooley, N. R. Crake, O. O. Dada, K. D. Diakoumakos, B. Dominguez-Fernandez, D. J. Earnshaw, U. C. Egbujor, D. W. Elmore, S. S. Etchin, M. R. Ewan, M. Fedurco, L. J. Fraser, K. V. Fuentes Fajardo, W. Scott Furey, D. George, K. J. Gietzen, C. P. Goddard, G. S. Golda, P. A. Granieri, D. E. Green, D. L. Gustafson, N. F. Hansen, K. Harnish, C. D. Haudenschild, N. I. Heyer, M. M. Hims, J. T. Ho, A. M. Horgan, K. Hoschler, S. Hurwitz, D. V. Ivanov, M. Q. Johnson, T. James, T. A. Huw Jones, G. D. Kang, T. H. Kerelska, A. D. Kersey, I. Khrebtukova, A. P. Kindwall, Z. Kingsbury, P. I. Kokko-Gonzales, A. Kumar, M. A. Laurent, C. T. Lawley, S. E. Lee, X. Lee, A. K. Liao, J. A. Loch, M. Lok, S. Luo, R. M. Mammen, J. W. Martin, P. G. McCauley, P. McNitt, P. Mehta, K. W. Moon, J. W. Mullens, T. Newington, Z. Ning, B. Ling Ng, S. M. Novo, M. J. O'Neill, M. A. Osborne, A. Osnowski, O. Ostadan, L. L. Paraschos, L. Pickering, A. C. Pike, A. C. Pike, D. Chris Pinkard, D. P. Pliskin, J. Podhasky, V. J. Quijano, C. Raczy, V. H. Rae, S. R. Rawlings, A. Chiva Rodriguez, P. M. Roe, J. Rogers, M. C. Rogert Bacigalupo, N. Romanov, A. Romieu, R. K. Roth, N. J. Rourke, S. T. Ruediger, E. Rusman, R. M. Sanches-Kuiper, M. R. Schenker, J. M. Seoane, R. J. Shaw, M. K. Shiver, S. W. Short, N. L. Sizto, J. P. Sluis, M. A. Smith, J. Ernest Sohna Sohna, E. J. Spence, K. Stevens, N. Sutton, L. Szajkowski, C. L. Tregidgo, G. Turcatti, S. Vandevondele, Y. Verhovsky, S. M. Virk, S. Wakelin, G. C. Walcott, J. Wang, G. J. Worsley, J. Yan, L. Yau, M. Zuerlein, J. Rogers, J. C. Mullikin, M. E. Hurles, N. J. McCooke, J. S. West, F. L. Oaks, P. L. Lundberg, D. Klenerman, R. Durbin, A. J. Smith, Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53-59 (2008).

58. S. J. Jones, J. Laskin, Y. Y. Li, O. L. Griffith, J. An, M. Bilenky, Y. S. Butterfield, T. Cezard, E. Chuah, R. Corbett, A. P. Fejes, M. Griffith, J. Yee, M. Martin, M. Mayo, N. Melnyk, R. D. Morin, T. J. Pugh, T. Severson, S. P. Shah, M. Sutcliffe, A. Tam, J. Terry, N. Thiessen, T. Thomson, R. Varhol, T. Zeng, Y. Zhao, R. A. Moore, D. G. Huntsman, I. Birol, M. Hirst, R. A. Holt, M. A. Marra, Evolution of an adenocarcinoma in response to selection by targeted kinase inhibitors. *Genome biology* **11**, R82 (2010).

59. C. International Cancer Genome, T. J. Hudson, W. Anderson, A. Artez, A. D. Barker, C. Bell, R. R. Bernabe, M. K. Bhan, F. Calvo, I. Eerola, D. S. Gerhard, A. Guttmacher, M. Guyer, F. M. Hemsley, J. L. Jennings, D. Kerr, P. Klatt, P. Kolar, J. Kusada, D. P. Lane, F. Laplace, L. Youyong, G. Nettekoven, B. Ozenberger, J. Peterson, T. S. Rao, J. Remacle, A. J. Schafer, T. Shibata, M. R. Stratton, J. G. Vockley, K. Watanabe, H. Yang, M. M. Yuen, B. M. Knoppers, M. Bobrow, A. Cambon-Thomsen, L. G. Dressler, S. O. Dyke, Y. Joly, K. Kato, K. L. Kennedy, P. Nicolas, M. J. Parker, E. Rial-Sebbag, C. M. Romeo-Casabona, K. M. Shaw, S. Wallace, G. L. Wiesner, N. Zeps, P. Lichter, A. V. Biankin, C. Chabannon, L. Chin, B. Clement, E. de Alava, F. Degos, M. L. Ferguson, P. Geary, D. N. Hayes, T. J. Hudson, A. L. Johns, A. Kasprzyk, H. Nakagawa, R. Penny, M. A. Piris, R. Sarin, A. Scarpa, T. Shibata, M. van de Vijver, P. A. Futreal, H. Aburatani, M. Bayes, D. D. Botwell, P. J. Campbell, X. Estivill, D. S. Gerhard, S. M. Grimmond, I. Gut, M. Hirst, C. Lopez-Otin, P. Majumder, M. Marra, J. D. McPherson, H. Nakagawa, Z. Ning, X. S. Puente, Y. Ruan, T. Shibata, M. R. Stratton, H. G. Stunnenberg, H. Swerdlow, V. E. Velculescu, R. K. Wilson, H. H. Xue, L. Yang, P. T. Spellman, G. D. Bader, P. C. Boutros, P. J. Campbell, P. Flicek, G. Getz, R. Guigo, G. Guo, D. Haussler, S. Heath, T. J. Hubbard, T. Jiang, S. M. Jones, Q. Li, N. Lopez-Bigas, R. Luo, L. Muthuswamy, B. F. Ouellette, J. V. Pearson, X. S. Puente, V. Quesada, B. J. Raphael, C. Sander, T. Shibata, T. P. Speed, L. D. Stein, J. M. Stuart, J. W. Teague, Y. Totoki, T. Tsunoda, A. Valencia, D. A. Wheeler, H. Wu, S. Zhao, G. Zhou, L. D. Stein, R. Guigo, T. J. Hubbard, Y. Joly, S. M. Jones, A. Kasprzyk, M. Lathrop, N. Lopez-Bigas, B. F. Ouellette, P. T. Spellman, J. W. Teague, G. Thomas, A. Valencia, T. Yoshida, K. L. Kennedy,

M. Axton, S. O. Dyke, P. A. Futreal, D. S. Gerhard, C. Gunter, M. Guyer, T. J. Hudson, J. D. McPherson, L. J. Miller, B. Ozenberger, K. M. Shaw, A. Kasprzyk, L. D. Stein, J. Zhang, S. A. Haider, J. Wang, C. K. Yung, A. Cros, Y. Liang, S. Gnaneshan, J. Guberman, J. Hsu, M. Bobrow, D. R. Chalmers, K. W. Hasel, Y. Joly, T. S. Kaan, K. L. Kennedy, B. M. Knoppers, W. W. Lowrance, T. Masui, P. Nicolas, E. Rial-Sebbag, L. L. Rodriguez, C. Vergely, T. Yoshida, S. M. Grimmond, A. V. Biankin, D. D. Bowtell, N. Cloonan, A. deFazio, J. R. Eshleman, D. Etemadmoghadam, B. B. Gardiner, J. G. Kench, A. Scarpa, R. L. Sutherland, M. A. Tempero, N. J. Waddell, P. J. Wilson, J. D. McPherson, S. Gallinger, M. S. Tsao, P. A. Shaw, G. M. Petersen, D. Mukhopadhyay, L. Chin, R. A. DePinho, S. Thayer, L. Muthuswamy, K. Shazand, T. Beck, M. Sam, L. Timms, V. Ballin, Y. Lu, J. Ji, X. Zhang, F. Chen, X. Hu, G. Zhou, Q. Yang, G. Tian, L. Zhang, X. Xing, X. Li, Z. Zhu, Y. Yu, J. Yu, H. Yang, M. Lathrop, J. Tost, P. Brennan, I. Holcatova, D. Zaridze, A. Brazma, L. Egevard, E. Prokhortchouk, R. E. Banks, M. Uhlen, A. Cambon-Thomsen, J. Viksna, F. Ponten, K. Skryabin, M. R. Stratton, P. A. Futreal, E. Birney, A. Borg, A. L. Borresen-Dale, C. Caldas, J. A. Foekens, S. Martin, J. S. Reis-Filho, A. L. Richardson, C. Sotiriou, H. G. Stunnenberg, G. Thoms, M. van de Vijver, L. van't Veer, F. Calvo, D. Birnbaum, H. Blanche, P. Boucher, S. Boyault, C. Chabannon, I. Gut, J. D. Masson-Jacquemier, M. Lathrop, I. Pauporte, X. Pivot, A. Vincent-Salomon, E. Tabone, C. Theillet, G. Thomas, J. Tost, I. Treilleux, F. Calvo, P. Bioulac-Sage, B. Clement, T. Decaens, F. Degos, D. Franco, I. Gut, M. Gut, S. Heath, M. Lathrop, D. Samuel, G. Thomas, J. Zucman-Rossi, P. Lichter, R. Eils, B. Brors, J. O. Korbel, A. Korshunov, P. Landgraf, H. Lehrach, S. Pfister, B. Radlwimmer, G. Reifenberger, M. D. Taylor, C. von Kalle, P. P. Majumder, R. Sarin, T. S. Rao, M. K. Bhan, A. Scarpa, P. Pederzoli, R. A. Lawlor, M. Delledonne, A. Bardelli, A. V. Biankin, S. M. Grimmond, T. Gress, D. Klimstra, G. Zamboni, T. Shibata, Y. Nakamura, H. Nakagawa, J. Kusada, T. Tsunoda, S. Miyano, H. Aburatani, K. Kato, A. Fujimoto, T. Yoshida, E. Campo, C. Lopez-Otin, X. Estivill, R. Guigo, S. de Sanjose, M. A. Piris, E. Montserrat, M. Gonzalez-Diaz, X. S. Puente, P. Jares, A. Valencia, H. Himmelbauer, V. Quesada, S. Bea, M. R. Stratton, P. A. Futreal, P. J. Campbell, A. Vincent-Salomon, A. L. Richardson, J. S. Reis-Filho, M. van de Vijver, G. Thomas, J. D. Masson-Jacquemier, S. Aparicio, A. Borg, A. L. Borresen-Dale, C. Caldas, J. A. Foekens, H. G. Stunnenberg, L. van't Veer, D. F. Easton, P. T. Spellman, S. Martin, A. D. Barker, L. Chin, F. S. Collins, C. C. Compton, M. L. Ferguson, D. S. Gerhard, G. Getz, C. Gunter, A. Guttmacher, M. Guyer, D. N. Hayes, E. S. Lander, B. Ozenberger, R. Penny, J. Peterson, C. Sander, K. M. Shaw, T. P. Speed, P. T. Spellman, J. G. Vockley, D. A. Wheeler, R. K. Wilson, T. J. Hudson, L. Chin, B. M. Knoppers, E. S. Lander, P. Lichter, L. D. Stein, M. R. Stratton, W. Anderson, A. D. Barker, C. Bell, M. Bobrow, W. Burke, F. S. Collins, C. C. Compton, R. A. DePinho, D. F. Easton, P. A. Futreal, D. S. Gerhard, A. R. Green, M. Guyer, S. R. Hamilton, T. J. Hubbard, O. P. Kallioniemi, K. L. Kennedy, T. J. Ley, E. T. Liu, Y. Lu, P. Majumder, M. Marra, B. Ozenberger, J. Peterson, A. J. Schafer, P. T. Spellman, H. G. Stunnenberg, B. J. Wainwright, R. K. Wilson, H. Yang, International network of cancer genome projects. *Nature* **464**, 993-998 (2010).

60. C. Curtis, S. P. Shah, S. F. Chin, G. Turashvili, O. M. Rueda, M. J. Dunning, D. Speed, A. G. Lynch, S. Samarajiwa, Y. Yuan, S. Graf, G. Ha, G. Haffari, A. Bashashati, R. Russell, S. McKinney, M. Group, A. Langerod, A. Green, E. Provenzano, G. Wishart, S. Pinder, P. Watson, F. Markowetz, L. Murphy, I. Ellis, A. Purushotham, A. L. Borresen-Dale, J. D. Brenton, S. Tavare, C. Caldas, S. Aparicio, The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-352 (2012).

61. N. Wan, D. Weinberg, T. Y. Liu, K. Niehaus, E. A. Ariazi, D. Delubac, A. Kannan, B. White, M. Bailey, M. Bertin, N. Boley, D. Bowen, J. Cregg, A. M. Drake, R. Ennis, S. Fransen, E. Gafni, L. Hansen, Y. Liu, G. L. Otte, J. Pecson, B. Rice, G. E. Sanderson, A. Sharma, J. St John, C. Tang, A. Tzou, L. Young, G. Putcha, I. S. Haque, Machine learning enables detection of early-stage colorectal cancer by whole-genome sequencing of plasma cell-free DNA. *BMC Cancer* **19**, 832 (2019).

62. I. Martincorena, A. Roshan, M. Gerstung, P. Ellis, P. Van Loo, S. McLaren, D. C. Wedge, A. Fullam, L. B. Alexandrov, J. M. Tubio, L. Stebbings, A. Menzies, S. Widaa, M. R. Stratton,

P. H. Jones, P. J. Campbell, Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science (New York, N.Y.)* **348**, 880-886 (2015).
63. A. P. Feinberg, The Key Role of Epigenetics in Human Disease Prevention and Mitigation. *The New England journal of medicine* **378**, 1323-1334 (2018).
64. Y. Long, X. Wang, D. T. Youmans, T. R. Cech, How do lncRNAs regulate transcription? *Science advances* **3**, 2110 (2017).
65. M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science (New York, N.Y.)* **324**, 930-935 (2009).
66. M. C. Liu, G. R. Oxnard, E. A. Klein, C. Swanton, M. V. Seiden, S. R. Cummings, F. Absalan, G. Alexander, B. Allen, H. Amini, A. M. Aravanis, S. Bagaria, L. Bazargan, J. F. Beausang, J. Berman, C. Betts, A. Blocker, J. Bredno, R. Calef, G. Cann, J. Carter, C. Chang, H. Chawla, X. Chen, T. C. Chien, D. Civello, K. Davydov, V. Demas, M. Desai, Z. Dong, S. Fayzullina, A. P. Fields, D. Filippova, P. Freese, E. T. Fung, S. Gnerre, S. Gross, M. Halks-Miller, M. P. Hall, A.-R. Hartman, C. Hou, E. Hubbell, N. Hunkapiller, K. Jagadeesh, A. Jamshidi, R. Jiang, B. Jung, T. Kim, R. D. Klausner, K. N. Kurtzman, M. Lee, W. Lin, J. Lipson, H. Liu, Q. Liu, M. Lopatin, T. Maddala, M. C. Maher, C. Melton, A. Mich, S. Nautiyal, J. Newman, J. Newman, V. Nicula, C. Nicolaou, O. Nikolic, W. Pan, S. Patel, S. A. Prins, R. Rava, N. Ronaghi, O. Sakarya, R. V. Satya, J. Schellenberger, E. Scott, A. J. Sehnert, R. Shaknovich, A. Shanmugam, K. C. Shashidhar, L. Shen, A. Shenoy, S. Shojaee, P. Singh, K. K. Steffen, S. Tang, J. M. Toung, A. Valouev, O. Venn, R. T. Williams, T. Wu, H. H. Xu, C. Yakym, X. Yang, J. Yecies, A. S. Yip, J. Youngren, J. Yue, J. Zhang, L. Zhang, L. Zhang, N. Zhang, C. Curtis, D. A. Berry, Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Annals of Oncology* **31**, 745-759 (2020).
67. S. Cristiano, A. Leal, J. Phallen, J. Fiksel, V. Adleff, D. C. Bruhm, S. O. Jensen, J. E. Medina, C. Hruban, J. R. White, D. N. Palsgrove, N. Niknafs, V. Anagnostou, P. Forde, J. Naidoo, K. Marrone, J. Brahmer, B. D. Woodward, H. Husain, K. L. van Rooijen, M. W. Orntoft, A. H. Madsen, C. J. H. van de Velde, M. Verheij, A. Cats, C. J. A. Punt, G. R. Vink, N. C. T. van Grieken, M. Koopman, R. J. A. Fijneman, J. S. Johansen, H. J. Nielsen, G. A. Meijer, C. L. Andersen, R. B. Scharpf, V. E. Velculescu, Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature* **570**, 385-389 (2019).
68. A. Carlsson, V. S. Nair, M. S. Luttgen, K. V. Keu, G. Horng, M. Vasanawala, A. Kolatkar, M. Jamali, A. H. Iagaru, W. Kuschner, B. W. Loo, Jr., J. B. Shrager, K. Bethel, C. K. Hoh, L. Bazhenova, J. Nieva, P. Kuhn, S. S. Gambhir, Circulating tumor microemboli diagnostics for patients with non-small-cell lung cancer. *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* **9**, 1111-1119 (2014).
69. S. H. Jalalian, M. Ramezani, S. A. Jalalian, K. Abnous, S. M. Taghdisi, Exosomes, new biomarkers in early cancer detection. *Anal Biochem* **571**, 1-13 (2019).
70. C. E. Gast, A. D. Silk, L. Zarour, L. Riegler, J. G. Burkhart, K. T. Gustafson, M. S. Parappilly, M. Roh-Johnson, J. R. Goodman, B. Olson, M. Schmidt, J. R. Swain, P. S. Davies, V. Shasthri, S. Iizuka, P. Flynn, S. Watson, J. Korkola, S. A. Courtneidge, J. M. Fischer, J. Jaboin, K. G. Billingsley, C. D. Lopez, J. Burchard, J. Gray, L. M. Coussens, B. C. Sheppard, M. H. Wong, Cell fusion potentiates tumor heterogeneity and reveals circulating hybrid cells that correlate with stage and survival. *Sci Adv* **4**, eaat7828 (2018).
71. A. Srivastava, D. J. Creek, Discovery and Validation of Clinical Biomarkers of Cancer: A Review Combining Metabolomics and Proteomics. *Proteomics* **19**, e1700448 (2019).
72. H. T. Tan, Y. H. Lee, M. C. Chung, Cancer proteomics. *Mass Spectrom Rev* **31**, 583-605 (2012).
73. M. Bhardwaj, K. Weigl, K. Tikk, T. Holland-Letz, P. Schrotz-King, C. H. Borchers, H. Brenner, Multiplex quantitation of 270 plasma protein markers to identify a signature for early detection of colorectal cancer. *European journal of cancer (Oxford, England : 1990)* **127**, 30-40 (2020).

74. M. Bhardwaj, A. Gies, S. Werner, P. Schrotz-King, H. Brenner, Blood-Based Protein Signatures for Early Detection of Colorectal Cancer: A Systematic Review. *Clinical and translational gastroenterology* **8**, e128 (2017).
75. J. D. Cohen, L. Li, Y. Wang, C. Thoburn, B. Afsari, L. Danilova, C. Douville, A. A. Javed, F. Wong, A. Mattox, R. H. Hruban, C. L. Wolfgang, M. G. Goggins, M. Dal Molin, T. L. Wang, R. Roden, A. P. Klein, J. Ptak, L. Dobbyn, J. Schaefer, N. Silliman, M. Popoli, J. T. Vogelstein, J. D. Browne, R. E. Schoen, R. E. Brand, J. Tie, P. Gibbs, H. L. Wong, A. S. Mansfield, J. Jen, S. M. Hanash, M. Falconi, P. J. Allen, S. Zhou, C. Bettegowda, L. A. Diaz, Jr., C. Tomasetti, K. W. Kinzler, B. Vogelstein, A. M. Lennon, N. Papadopoulos, Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science (New York, N.Y.)* **359**, 926-930 (2018).
76. V. Gopalakrishnan, C. N. Spencer, L. Nezi, A. Reuben, M. C. Andrews, T. V. Karpinets, P. A. Prieto, D. Vicente, K. Hoffman, S. C. Wei, A. P. Cogdill, L. Zhao, C. W. Hudgens, D. S. Hutchinson, T. Manzo, M. Petaccia de Macedo, T. Cotechini, T. Kumar, W. S. Chen, S. M. Reddy, R. Szczepaniak Sloane, J. Galloway-Pena, H. Jiang, P. L. Chen, E. J. Shpall, K. Rezvani, A. M. Alousi, R. F. Chemaly, S. Shelburne, L. M. Vence, P. C. Okhuysen, V. B. Jensen, A. G. Swennes, F. McAllister, E. Marcelo Riquelme Sanchez, Y. Zhang, E. Le Chatelier, L. Zitvogel, N. Pons, J. L. Austin-Breneman, L. E. Haydu, E. M. Burton, J. M. Gardner, E. Sirmans, J. Hu, A. J. Lazar, T. Tsujikawa, A. Diab, H. Tawbi, I. C. Glitza, W. J. Hwu, S. P. Patel, S. E. Woodman, R. N. Amaria, M. A. Davies, J. E. Gershenwald, P. Hwu, J. E. Lee, J. Zhang, L. M. Coussens, Z. A. Cooper, P. A. Futreal, C. R. Daniel, N. J. Ajami, J. F. Petrosino, M. T. Tetzlaff, P. Sharma, J. P. Allison, R. R. Jenq, J. A. Wargo, Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science (New York, N.Y.)* **359**, 97-103 (2018).
77. V. Matson, J. Fessler, R. Bao, T. Chongsuwat, Y. Zha, M. L. Alegre, J. J. Luke, T. F. Gajewski, The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science (New York, N.Y.)* **359**, 104-108 (2018).
78. B. Routy, E. Le Chatelier, L. Derosa, C. P. M. Duong, M. T. Alou, R. Daillere, A. Fluckiger, M. Messaoudene, C. Rauber, M. P. Roberti, M. Fidelle, C. Flament, V. Poirier-Colame, P. Opolon, C. Klein, K. Iribarren, L. Mondragon, N. Jacquelot, B. Qu, G. Ferrere, C. Clemenson, L. Mezquita, J. R. Masip, C. Naltet, S. Brosseau, C. Kaderbhai, C. Richard, H. Rizvi, F. Levenez, N. Galleron, B. Quinquis, N. Pons, B. Ryffel, V. Minard-Colin, P. Gonin, J. C. Soria, E. Deutsch, Y. Loriot, F. Ghiringhelli, G. Zalcman, F. Goldwasser, B. Escudier, M. D. Hellmann, A. Eggermont, D. Raoult, L. Albiges, G. Kroemer, L. Zitvogel, Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science (New York, N.Y.)* **359**, 91-97 (2018).
79. B. A. Helmink, M. A. W. Khan, A. Hermann, V. Gopalakrishnan, J. A. Wargo, The microbiome, cancer, and cancer therapy. *Nature medicine* **25**, 377-388 (2019).
80. S. U. Springer, C. H. Chen, M. D. C. Rodriguez Pena, L. Li, C. Douville, Y. Wang, J. D. Cohen, D. Taheri, N. Silliman, J. Schaefer, J. Ptak, L. Dobbyn, M. Papoli, I. Kinde, B. Afsari, A. C. Tregnago, S. M. Bezerra, C. VandenBussche, K. Fujita, D. Ertoy, I. W. Cunha, L. Yu, T. J. Bivalacqua, A. P. Grollman, L. A. Diaz, R. Karchin, L. Danilova, C. Y. Huang, C. T. Shun, R. J. Turesky, B. H. Yun, T. A. Rosenquist, Y. S. Pu, R. H. Hruban, C. Tomasetti, N. Papadopoulos, K. W. Kinzler, B. Vogelstein, K. G. Dickman, G. J. Netto, Non-invasive detection of urothelial cancer through the analysis of driver gene mutations and aneuploidy. *eLife* **7**, (2018).
81. Y. Wang, L. Li, C. Douville, J. D. Cohen, T. T. Yen, I. Kinde, K. Sundfelt, S. K. Kjaer, R. H. Hruban, I. M. Shih, T. L. Wang, R. J. Kurman, S. Springer, J. Ptak, M. Popoli, J. Schaefer, N. Silliman, L. Dobbyn, E. J. Tanner, A. Angarita, M. Lycke, K. Jochumsen, B. Afsari, L. Danilova, D. A. Levine, K. Jardon, X. Zeng, J. Arseneau, L. Fu, L. A. Diaz, Jr., R. Karchin, C. Tomasetti, K. W. Kinzler, B. Vogelstein, A. N. Fader, L. Gilbert, N. Papadopoulos, Evaluation of liquid from the Papanicolaou test and other liquid biopsies for the detection of endometrial and ovarian cancers. *Science translational medicine* **10**, (2018).
82. Z. Hu, J. Ding, Z. Ma, R. Sun, J. A. Seoane, J. Scott Shaffer, C. J. Suarez, A. S. Berghoff, C. Cremolini, A. Falcone, F. Loupakis, P. Birner, M. Preusser, H. J. Lenz, C. Curtis,

Quantitative evidence for early metastatic seeding in colorectal cancer. *Nature genetics* **51**, 1113-1122 (2019).

83. A. M. Lennon, A. H. Buchanan, I. Kinde, A. Warren, A. Honushefsky, A. T. Cohain, D. H. Ledbetter, F. Sanfilippo, K. Sheridan, D. Rosica, C. S. Adonizio, H. J. Hwang, K. Lahouel, J. D. Cohen, C. Douville, A. A. Patel, L. N. Hagmann, D. D. Rolston, N. Malani, S. Zhou, C. Bettegowda, D. L. Diehl, B. Urban, C. D. Still, L. Kann, J. I. Woods, Z. M. Salvati, J. Vadakara, R. Leeming, P. Bhattacharya, C. Walter, A. Parker, C. Lengauer, A. Klein, C. Tomasetti, E. K. Fishman, R. H. Hruban, K. W. Kinzler, B. Vogelstein, N. Papadopoulos, Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. *Science (New York, N.Y.)*, (2020).
84. T. F. Imperiale, D. F. Ransohoff, S. H. Itzkowitz, Multitarget stool DNA testing for colorectal-cancer screening. *The New England journal of medicine* **371**, 187-188 (2014).
85. V. Baeriswyl, G. Christofori, The angiogenic switch in carcinogenesis. *Seminars in cancer biology* **19**, 329-337 (2009).
86. X. He, M. A. Ertürk, A. Grant, X. Wu, R. L. Lagore, L. DelaBarre, Y. Eryaman, G. Adriany, E. J. Auerbach, P. F. Van de Moortele, K. Uğurbil, G. J. Metzger, First in-vivo human imaging at 10.5T: Imaging the body at 447 MHz. *Magnetic resonance in medicine* **84**, 289-303 (2020).
87. A. P. Soleimany, S. N. Bhatia, Activity-Based Diagnostics: An Emerging Paradigm for Disease Detection and Monitoring. *Trends in molecular medicine* **26**, 450-468 (2020).
88. S. Slomovic, K. Pardee, J. J. Collins, Synthetic biology devices for in vitro and in vivo diagnostics. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 14429-14435 (2015).
89. C. N. Loynachan, A. P. Soleimany, J. S. Dudani, Y. Lin, A. Najer, A. Bekdemir, Q. Chen, S. N. Bhatia, M. M. Stevens, Renal clearable catalytic gold nanoclusters for in vivo disease monitoring. *Nature nanotechnology* **14**, 883-890 (2019).
90. T. Danino, A. Prindle, G. A. Kwong, M. Skalak, H. Li, K. Allen, J. Hasty, S. N. Bhatia, Programmable probiotics for detection of cancer in urine. *Science translational medicine* **7**, 289ra284 (2015).
91. A. Aalipour, H. Y. Chuang, S. Murty, A. L. D'Souza, S. M. Park, G. S. Gulati, C. B. Patel, C. Beinat, F. Simonetta, I. Martinic, G. Gowrishankar, E. R. Robinson, E. Aalipour, Z. Zhian, S. S. Gambhir, Engineered immune cells as highly sensitive cancer diagnostics. *Nature biotechnology* **37**, 531-539 (2019).
92. A. F. Sarioglu, N. Aceto, N. Kojic, M. C. Donaldson, M. Zeinali, B. Hamza, A. Engstrom, H. Zhu, T. K. Sundaresan, D. T. Miyamoto, X. Luo, A. Bardia, B. S. Wittner, S. Ramaswamy, T. Shioda, D. T. Ting, S. L. Stott, R. Kapur, S. Maheswaran, D. A. Haber, M. Toner, A microfluidic device for label-free, physical capture of circulating tumor cell clusters. *Nature methods* **12**, 685-691 (2015).
93. P. Zhang, X. Zhou, M. He, Y. Shang, A. L. Tetlow, A. K. Godwin, Y. Zeng, Ultrasensitive detection of circulating exosomes with a 3D-nanopatterned microfluidic chip. *Nature biomedical engineering* **3**, 438-451 (2019).
94. S. S. Gambhir, T. J. Ge, O. Vermesh, R. Spitler, Toward achieving precision health. *Science translational medicine* **10**, (2018).
95. J. C. Gore, H. C. Manning, C. C. Quarles, K. W. Waddell, T. E. Yankeelov, Magnetic resonance in the era of molecular imaging of cancer. *Magnetic resonance imaging* **29**, 587-600 (2011).
96. J. V. Frangioni, New technologies for human cancer imaging. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **26**, 4012-4021 (2008).
97. S. Surti, A. Kuhn, M. E. Werner, A. E. Perkins, J. Kolthammer, J. S. Karp, Performance of Philips Gemini TF PET/CT scanner with special consideration for its time-of-flight imaging capabilities. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* **48**, 471-480 (2007).
98. K. Golman, R. I. Zandt, M. Lerche, R. Pehrson, J. H. Ardenkjaer-Larsen, Metabolic imaging by hyperpolarized 13C magnetic resonance imaging for in vivo tumor diagnosis. *Cancer research* **66**, 10855-10860 (2006).

99. H. U. Kauczor, A. M. Baird, T. G. Blum, L. Bonomo, C. Bostantzoglou, O. Burghuber, B. Cepicka, A. Comanescu, S. Couraud, A. Devaraj, V. Jespersen, S. Morozov, I. N. Agmon, N. Peled, P. Powell, H. Prosch, S. Ravara, J. Rawlinson, M. P. Revel, M. Silva, A. Snoeckx, B. van Ginneken, J. P. van Meerbeeck, C. Vardavas, O. von Stackelberg, M. Gaga, R. European Society of, S. the European Respiratory, ESR/ERS statement paper on lung cancer screening. *European Radiology*, (2020).
100. H. U. Ahmed, A. El-Shater Bosaily, L. C. Brown, R. Gabe, R. Kaplan, M. K. Parmar, Y. Collaco-Moraes, K. Ward, R. G. Hindley, A. Freeman, A. P. Kirkham, R. Oldroyd, C. Parker, M. Emberton, P. s. group, Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet (London, England)* **389**, 815-822 (2017).
101. J. Katzen, K. Dodelzon, A review of computer aided detection in mammography. *Clinical imaging* **52**, 305-309 (2018).
102. H. J. Yoon, J. Kang, H. Park, I. Sohn, S. H. Lee, H. Y. Lee, Deciphering the tumor microenvironment through radiomics in non-small cell lung cancer: Correlation with immune profiles. *PloS one* **15**, e0231227 (2020).
103. S. M. McKinney, M. Sieniek, V. Godbole, J. Godwin, N. Antropova, H. Ashrafian, T. Back, M. Chesus, G. C. Corrado, A. Darzi, M. Etemadi, F. Garcia-Vicente, F. J. Gilbert, M. Halling-Brown, D. Hassabis, S. Jansen, A. Karthikesalingam, C. J. Kelly, D. King, J. R. Ledsam, D. Melnick, H. Mostofi, L. Peng, J. J. Reicher, B. Romera-Paredes, R. Sidebottom, M. Suleyman, D. Tse, K. C. Young, J. De Fauw, S. Shetty, International evaluation of an AI system for breast cancer screening. *Nature* **577**, 89-94 (2020).
104. M. Bahl, R. Barzilay, A. B. Yedidia, N. J. Locascio, L. Yu, C. D. Lehman, High-Risk Breast Lesions: A Machine Learning Model to Predict Pathologic Upgrade and Reduce Unnecessary Surgical Excision. *Radiology* **286**, 810-818 (2018).
105. B. Haibe-Kains, G. A. Adam, A. Hosny, F. Khodakarami, L. Waldron, B. Wang, C. McIntosh, A. Goldenberg, A. Kundaje, C. S. Greene, T. Broderick, M. M. Hoffman, J. T. Leek, K. Korthauer, W. Huber, A. Brazma, J. Pineau, R. Tibshirani, T. Hastie, J. P. A. Ioannidis, J. Quackenbush, H. Aerts, Transparency and reproducibility in artificial intelligence. *Nature* **586**, E14-e16 (2020).
106. I. Steinberg, D. M. Huland, O. Vermesh, H. E. Frostig, W. S. Tummers, S. S. Gambhir, Photoacoustic clinical imaging. *Photoacoustics* **14**, 77-98 (2019).
107. W. B. Nagengast, E. Hartmans, P. B. Garcia-Allende, F. T. M. Peters, M. D. Linssen, M. Koch, M. Koller, J. J. J. Tjalma, A. Karrenbeld, A. Jorritsma-Smit, J. H. Kleibeuker, G. M. van Dam, V. Ntziachristos, Near-infrared fluorescence molecular endoscopy detects dysplastic oesophageal lesions using topical and systemic tracer of vascular endothelial growth factor A. *Gut* **68**, 7-10 (2019).
108. J. Burggraaf, I. M. Kamerling, P. B. Gordon, L. Schrier, M. L. de Kam, A. J. Kales, R. Bendiksen, B. Indrevoll, R. M. Bjerke, S. A. Moestue, S. Yazdanfar, A. M. Langers, M. Swaerd-Nordmo, G. Torheim, M. V. Warren, H. Morreau, P. W. Voorneveld, T. Buckle, F. W. van Leeuwen, L. I. Odegardstuen, G. T. Dalsgaard, A. Healey, J. C. Hardwick, Detection of colorectal polyps in humans using an intravenously administered fluorescent peptide targeted against c-Met. *Nature medicine* **21**, 955-961 (2015).
109. R. Woitek, F. A. Gallagher, The use of hyperpolarised (13)C-MRI in clinical body imaging to probe cancer metabolism. *British journal of cancer*, (2021).
110. K. Bera, K. A. Schalper, D. L. Rimm, V. Velcheti, A. Madabhushi, Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nature reviews. Clinical oncology* **16**, 703-715 (2019).
111. M. K. K. Niazi, A. V. Parwani, M. N. Gurcan, Digital pathology and artificial intelligence. *The Lancet. Oncology* **20**, e253-e261 (2019).
112. T. National Lung Screening Trial Research, D. R. Aberle, A. M. Adams, C. D. Berg, W. C. Black, J. D. Clapp, R. M. Fagerstrom, I. F. Gareen, C. Gatsonis, P. M. Marcus, J. D. Sicks, Reduced lung-cancer mortality with low-dose computed tomographic screening. *The New England journal of medicine* **365**, 395-409 (2011).

113. H. J. de Koning, C. M. van der Aalst, P. A. de Jong, E. T. Scholten, K. Nackaerts, M. A. Heuvelmans, J. J. Lammers, C. Weenink, U. Yousaf-Khan, N. Horeweg, S. van 't Westeinde, M. Prokop, W. P. Mali, F. A. A. Mohamed Hoesein, P. M. A. van Ooijen, J. Aerts, M. A. den Bakker, E. Thunnissen, J. Verschakelen, R. Vliegenthart, J. E. Walter, K. Ten Haaf, H. J. M. Groen, M. Oudkerk, Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. *The New England journal of medicine* **382**, 503-513 (2020).
114. U. Menon, A. Gentry-Maharaj, R. Hallett, A. Ryan, M. Burnell, A. Sharma, S. Lewis, S. Davies, S. Philpott, A. Lopes, K. Godfrey, D. Oram, J. Herod, K. Williamson, M. W. Seif, I. Scott, T. Mould, R. Woolas, J. Murdoch, S. Dobbs, N. N. Amso, S. Leeson, D. Cruickshank, A. McGuire, S. Campbell, L. Fallowfield, N. Singh, A. Dawnay, S. J. Skates, M. Parmar, I. Jacobs, Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *The Lancet. Oncology* **10**, 327-340 (2009).
115. J. Hugosson, M. J. Roobol, M. Mansson, T. L. J. Tammela, M. Zappa, V. Nelen, M. Kwiatkowski, M. Lujan, S. V. Carlsson, K. M. Talala, H. Lilja, L. J. Denis, F. Recker, A. Paez, D. Puliti, A. Villers, X. Rebillard, T. P. Kilpelainen, U. H. Stenman, R. A. Godtman, K. Stinesen Kollberg, S. M. Moss, P. Kujala, K. Taari, A. Huber, T. van der Kwast, E. A. Heijnsdijk, C. Bangma, H. J. De Koning, F. H. Schroder, A. Auvinen, E. investigators, A 16-yr Follow-up of the European Randomized study of Screening for Prostate Cancer. *European urology* **76**, 43-51 (2019).
116. L. Highfield, S. S. Rajan, M. A. Valerio, G. Walton, M. E. Fernandez, L. K. Bartholomew, A non-randomized controlled stepped wedge trial to evaluate the effectiveness of a multi-level mammography intervention in improving appointment adherence in underserved women. *Implementation science : IS* **10**, 143 (2015).
117. J. Cuzick, F. H. Cafferty, R. Edwards, H. Møller, S. W. Duffy, Surrogate endpoints for cancer screening trials: general principles and an illustration using the UK Flexible Sigmoidoscopy Screening Trial. *Journal of medical screening* **14**, 178-185 (2007).
118. G. Poste, Bring on the biomarkers. *Nature* **469**, 156-157 (2011).
119. S. A. Mulherin, W. C. Miller, Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. *Annals of internal medicine* **137**, 598-602 (2002).
120. B. S. Kramer, The science of early detection. *Urologic oncology* **22**, 344-347 (2004).
121. T. M. Andersson, M. J. Rutherford, K. Humphreys, Assessment of lead-time bias in estimates of relative survival for breast cancer. *Cancer epidemiology* **46**, 50-56 (2017).
122. National Cancer Institute, Early Detection Research Network. (2021) Available at: Accessed: 23 March 2021
123. National Cancer Institute, Consortium for Molecular and Cellular Characterization of Screen-Detected Lesions. (2021) Available at: https://mcl.nci.nih.gov/ Accessed: 23 March 2021
124. Cancer Research UK, Early Detection and Diagnosis Research Committee. (2021) Available at: https://www.cancerresearchuk.org/funding-for-researchers/applying-for-funding/funding-committees/early-detection-and-diagnosis-committee Accessed: 23 March 2021
125. Cancer Research UK, Early Detection and Diagnosis Roadmap. (2021) Available at: https://www.cancerresearchuk.org/funding-for-researchers/research-opportunities-in-early-detection-and-diagnosis/early-detection-and-diagnosis-roadmap Accessed: 23 March 2021
126. W. N. Hait, P. F. Lebowitz, Disease Interception: Myths, Mountains, and Mole Hills. *Cancer prevention research (Philadelphia, Pa.)* **9**, 635-637 (2016).
127. D. Schenkel, S. Nambi, R. Blicker, C. Lin, Liquid Biopsy: Early Detection of a Huge Investment Opportunity. (2020) Available at: https://www.cowen.com/insights/liquid-biopsy-early-detection-of-a-huge-investment-opportunity/ Accessed: 23 March 2021
128. Grand View Research, Cancer Diagnostics Market Worth $249.6 Billion By 2026. (2019) Available at: https://www.grandviewresearch.com/press-release/global-cancer-diagnostics-market Accessed: 23 March 2021
129. Office for National statistics UK, Cancer survival in England: adult, stage at diagnosis and childhood - patients followed up to 2018. (2019) Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddi

seases/bulletins/cancersurvivalinengland/stageatdiagnosisandchildhoodpatientsfollowedupto2018 Accessed: 23 March 2021