

# Radiomics Improves Cancer Screening and Early Detection **A C**

Robert J. Gillies<sup>1,2</sup> and Matthew B. Schabath<sup>3,4</sup>



## ABSTRACT

Imaging is a key technology in the early detection of cancers, including X-ray mammography, low-dose CT for lung cancer, or optical imaging for skin, esophageal, or colorectal cancers. Historically, imaging information in early detection schema was assessed qualitatively. However, the last decade has seen increased development of computerized tools that convert images into quantitative mineable data (radiomics), and their subsequent analyses with artificial intelligence (AI). These tools are improving diagnostic accuracy of early lesions to define risk and classify malignant/aggressive from benign/indolent disease. The first section of this review will briefly describe the various imaging modalities and their use as primary or secondary screens in an early detection pipeline.

The second section will describe specific use cases to illustrate the breadth of imaging modalities as well as the benefits of quantitative image analytics. These will include optical (skin cancer), X-ray CT (pancreatic and lung cancer), X-ray mammography (breast cancer), multiparametric MRI (breast and prostate cancer), PET (pancreatic cancer), and ultrasound elastography (liver cancer). Finally, we will discuss the inexorable improvements in radiomics to build more robust classifier models and the significant limitations to this development, including access to well-annotated databases, and biological descriptors of the imaged feature data.

See all articles in this *CEBP Focus* section, “NCI Early Detection Research Network: Making Cancer Detection Possible.”

## Introduction

As discussed elsewhere in this special issue, there are compelling medical and social needs for detecting cancers at an early stage while they are still localized and can be treated with curative intent. Ideally, the incipient cancers are detected in an asymptomatic individual because, in many cases, the onset of symptoms is associated with late-stage incurable disease. However, it is also abundantly clear that false positives, overdiagnosis, and overtreatment are serious issues that have to be solved before early detection paradigms can be fully successful (1). Consequently, early detection paradigms are necessarily multistep processes with a goal to optimize different tradeoffs between sensitivity and specificity of the diagnosis at each step. Historically, medical imaging has been a key component of early detection pipelines for all cancers (Tables 1 and 2).

In this review, we will present the case that quantitative analyses of these images with “radiomics” provides superior clinical utility to optimize diagnosis and risk assessment compared with qualitative interpretations of images. Furthermore, the resulting predictive machine learned models can be tuned to optimize sensitivity or specificity, depending upon whether the imaging test is a primary, or a secondary filter, and depending on the biology of the incipient lesion (2). Primary screens should have very high sensitivity, with increasing specificity in subsequent (secondary) screens.

## Radiomics

Radiomics refers to the conversion of images into structured, mineable data and the subsequent use of these data for prediction, diagnosis, prognosis, and/or longitudinal monitoring. The entire enterprise is predicated on the hypothesis that “Images are Data” and that they reflect the underlying pathobiology of the region of interest (ROI; ref. 3). Radiomics has been an explosive field in the last decade of its existence and has been thoroughly and recently reviewed elsewhere (4–6). The conventional radiomic analytic pipeline proceeds as shown in Fig. 1, wherein image-based features are extracted from the lesion and/or its surrounding tissues (i.e., ROIs). These features can be scored semantically by an experienced radiologist, or the features can be computer calculated from a hand-crafted set of features that number in the many hundreds to describe the size, shape, and textures within (and around) the lesion, or the features can be created by computed deep-learning (DL) algorithms. In some cases, “delta” features may also be computed by comparing individual feature values obtained at different times across screening intervals, during active surveillance, or during therapy (7).

The calculated features are then commonly analyzed with conventional machine learning approaches to develop models predictive of the dependent variable(s) of interest (8). In machine learning, the easiest outcomes to model are binary, and these are often most relevant to early detection paradigms, for example, cancer versus not cancer, aggressive versus indolent, etc. However, machine learning algorithms are also capable of modeling more complex continuous outcomes such as progression-free survival and time to recurrence, although these generally require much larger training sets. To avoid overfitting, the radiomic feature set must be reduced to a manageable number and this is done by sequentially removing unstable and redundant features, and then prioritizing those that are the most informative of outcome of interest. Whenever possible, radiomic models are combined with orthogonal data such as those provided by clinical covariates such as demographics, patient risk factors, serum analytes, or genomics. Such models that include both radiomics and clinical covariates generally improves predictive capabilities.

In conventional radiomics, accurate segmentation and choice of user-defined features are critical to success. Alternatively, DL

<sup>1</sup>Department of Cancer Physiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida. <sup>2</sup>Department of Radiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida. <sup>3</sup>Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida. <sup>4</sup>Department of Thoracic Oncology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida.

**Corresponding Author:** Robert J. Gillies, H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Drive, SRB 24000E, Tampa, FL 33612. Phone: 813-745-8355; Fax: 813-979-7265; E-mail: Robert.Gillies@moffitt.org

Cancer Epidemiol Biomarkers Prev 2020;29:2556–67

doi: 10.1158/1055-9965.EPI-20-0075

©2020 American Association for Cancer Research.

**Table 1.** Imaging modalities in early detection paradigms.

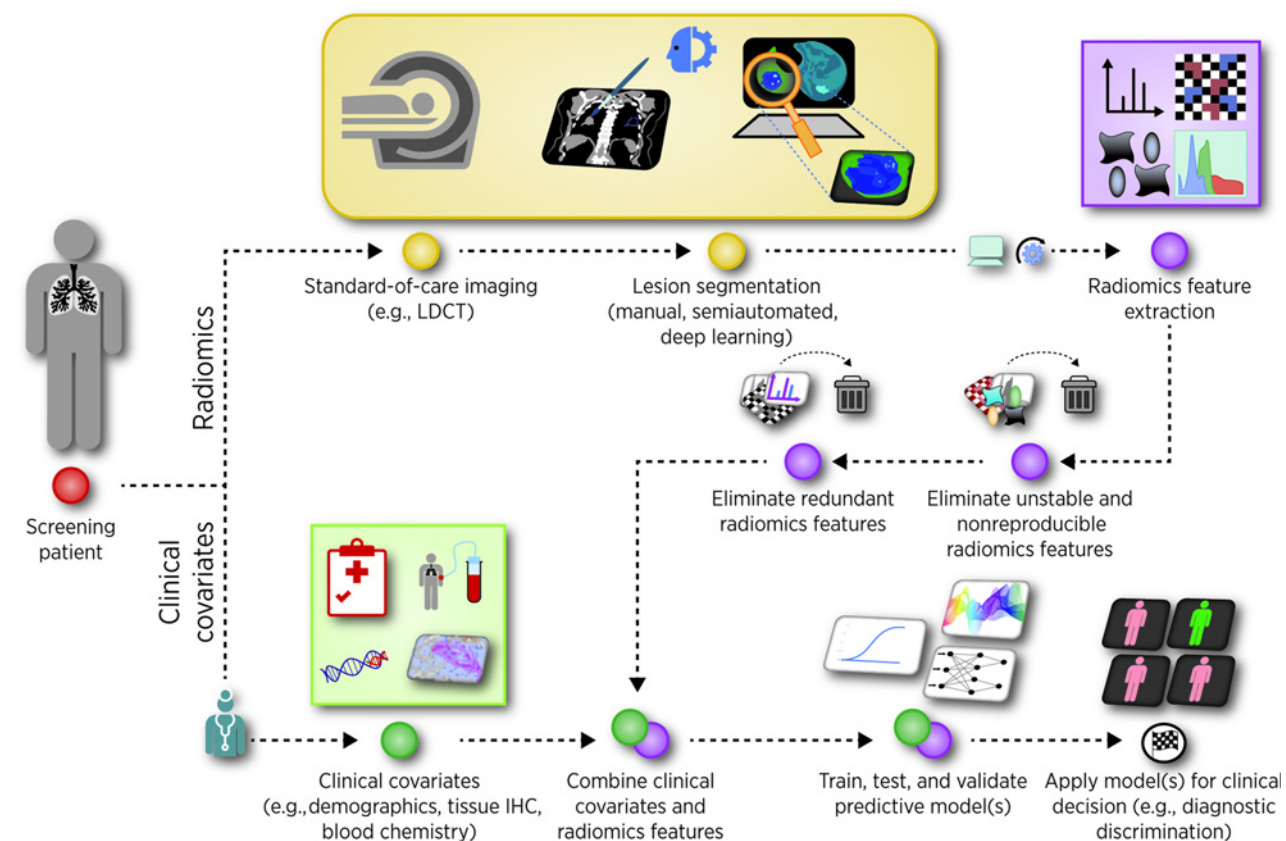
| Cancer     | Primary screen(s)       | Secondary screen(s)   |
|------------|-------------------------|-----------------------|
| Lung       | LDCT/DxCT               | Endoscopy, PET, Bx    |
| Breast     | Mammography/digital     | US, mpMRI, Bx         |
| Pancreatic | Serum, pancreatic juice | CT radiomics, MRI, Bx |
| Skin       | Optical                 | Bx                    |
| GI         | Optical (OCT)           | Bx                    |
| Liver/HCC  | CT/MRI                  | US, MR elastography   |
| Cervical   | Optical                 | Bx                    |

Abbreviations: Bx, biopsy; DxCT, diagnostic CT; GI, gastrointestinal; HCC, hepatocellular carcinoma; LDCT, low-dose CT; mpMRI, multiparametric MRI; OCT, optical coherence tomography; US, ultrasound.

approaches can mitigate these limitations as the algorithms self-train to identify the regions of the image that are most highly correlated to outcome (usually binary), and to identify the characteristics of that region that informed that decision through multiple layers of learning. For medical images, convolutional neural networks (CNN) are the most common models as they can accept two-dimensional or three-dimensional images as input. DL has only

recently been applied to radiomics but it has already proven valuable in differential diagnosis (9). The major limitation affecting the use of CNNs in diagnostic imaging is the availability of training data that are adequate both in terms of quality and quantity. There are approaches to solve this by limiting the size of the input images, artificially increasing the training set size via “transfer learning” that trains on an unrelated set of images or by image transformation using geometric augmentation (4).

Regardless of which radiomic approach and pipeline is utilized, the greatest single limitation is access to large, high-quality datasets that are well annotated with respect to the relevant outcome(s). The importance of this last criterion cannot be overstated. For many radiomic studies in the cancer care continuum, outcomes such as, *inter alia*, recurrence, progression-free survival, or time to recurrence are rarely captured in readily accessible structured formats. Hence, these endpoints need to be captured manually through chart review that is laborious and requires relevant expertise to extract accurately. In an early detection paradigm, however, the important outcome (cancer vs. noncancer; indolent vs. aggressive) is often, or can be, captured in structured formats, mitigating some of the limitations associated with data curation. Nonetheless, ongoing access to large and evolving datasets remains



**Figure 1.** The radiomics pipeline. Patient standard-of-care imaging and data (as available) are obtained prospectively or retrospectively. A radiologist or imaging scientist identifies the ROIs (e.g., pulmonary nodule), which are then segmented manually or by using semiautomatic algorithms or through DL algorithms. After the ROI is segmented, radiomic features are extracted (purple). Unstable, nonreproducible, and correlated radiomic features are eliminated. The remaining radiomics features are combined with relevant clinical covariates (green) and predictive model building approaches are applied, which can be applied for clinical decision. Figure 1 is modified from Tunali and colleagues (113). LDCT, low-dose CT.

Downloaded from <http://aacrjournals.org/cebp/article-pdf/29/12/2556/1948353/2556.pdf> by guest on 28 August 2022

**Table 2.** Strengths and limitations of imaging modalities in early detection.

| Modality | Submodality        | Use case     | Ease of use | Sensitivity and specificity | Cost | Other strengths | Other limitations |
|----------|--------------------|--------------|-------------|-----------------------------|------|-----------------|-------------------|
| Optical  | White light        | Skin cancer  | ****        | ***                         | *    |                 |                   |
|          |                    | GI endo      | ****        | ****                        | **** |                 |                   |
|          | OCT                | GI endo      | ****        | ***                         | **   | Depth           | Uncommon          |
|          |                    | Fluorescence | GI endo     | ***                         | **** |                 |                   |
| X-ray    | Mammography        | Breast       | **          | **                          | **   |                 | Evidence          |
|          |                    | Breast       | ***         | ***                         | **   | Radiation       |                   |
|          | CT                 | Lung, PDAC   | ***         | ***                         | **   | Radiation       |                   |
| US       | CEUS               | CRC, BrCa    | ***         | ***                         | **   |                 | Subjective        |
|          | Elastography       | Liver        | ****        | ***                         | *    |                 | Uncommon          |
| MRI      | Diffusion          | Whole body   | ***         | ****                        | **** | No radiation    | Uncommon          |
|          | DCE                | Breast       | ***         | ****                        | **** |                 |                   |
|          | Elastography       | Liver        | ***         | ****                        | **** |                 |                   |
| Nuclear  | $\alpha\beta6$ PET | PDAC         | ****        | ****                        | **** |                 | Uncommon          |
|          | FDG PET            | GI           | **          | ****                        | **** |                 |                   |

Note: The scoring chart ranges from \* to \*\*\*\*. Ease of use is scored from difficult to easy to use. Sensitivity and specificity are scored from low to high sensitivity and specificity. Cost is scored from low to high cost. Abbreviations: BrCa, breast cancer; CEUS, contrast-enhanced ultrasound; CRC, colorectal cancer; DCE, dynamic contrast-enhanced; GI, gastrointestinal; OCT, optical coherence tomography; PDAC, pancreatic ductal adenocarcinoma; US, ultrasound.

the biggest single limitation to the widespread development of radiomic approaches for prediction, prognosis, and monitoring. One approach to solve this problem is the application of “distributed learning” wherein algorithms, and not data, are shared between institutions for training and eventual testing (10, 11).

A direct consequence of the paucity of large multi-institutional and complex datasets is reduced generalizability and portability of the developed algorithms. In the discovery phase of Radiomics, single institution moderately sized datasets were acceptable. However, it is axiomatic that clinical utility of such approaches will require that the predictive accuracy of the algorithms be independent of site, scanner, or reconstruction parameters. Hence, there is an absolute requirement for models to be trained on complex multisite, multivendor, and multiplatform datasets. The specific context of use (COU), required for every biomarker, will then be defined by the complexity of the associated test and validation sets (12).

Another serious limitation to the use of radiomic algorithms, whether conventional or deep learned, is identification of the biological underpinnings of the most “informative” (i.e., predictive) radiomic features. For example, recognizing that it will be difficult to make treatment decisions based on a highly informative “Grey Length Co-occurrence Matrix” (GLCM) feature, it is incumbent upon the radiomics community to identify the biologically relevant drivers of the informative radiomics models. Examples in lung cancer include radiomic predictors of *EGFR* mutational status (13), immune infiltration (14), or PD-L1 status (15), which could conceivably then be used to trigger treatment decisions.

### Imaging Modalities

As shown in **Tables 1** and **2**, all imaging modalities contribute to the mission of early detection, including optical, X-rays (CT and mammography), ultrasound (US), MRI, and PET.

#### Optical

Optical approaches have the benefit of being relatively inexpensive and widely available. The major limitation to optical techniques is penetration depth that is limited to 1 to 2 mm. Although this can

be overcome with optical coherence (OCT) or optoacoustic (OA) imaging, these techniques are not yet widely available and have yet to be tested in an early detection pipeline. Similarly, molecular imaging with targeted fluorophores or nanoprobe (16) has great promise, but these are still very much in the preclinical or early clinical research setting. For example, OCT is being tested for their ability to improve detection of colorectal cancer using sigmoidoscopy (17). Currently, the most widespread use of optical imaging is the classification of skin lesions as being worrisome enough to send to biopsy and surgery. This is being greatly improved with artificial intelligence (AI), which will be discussed in the subsequent section.

#### US

US has the decided advantages of being able to image deep body tissues in real time and be very widely available, but it has been plagued in early detection schema by being too operator dependent. This is being aggressively solved using a family of techniques called “elastography,” which can quantitatively image tissue stiffness that is well known to be increased with hyperplasia. Elastography can potentially be used to detect lymph node metastases (18) or incipient breast lesions, but the current leading application is quantification of liver fibrosis, which will be discussed in the subsequent section.

#### X-ray

High-energy photons have superior tissue penetration and can be absorbed by electron-dense materials, such as calcium deposits, or contrast agents. These are also absorbed by densities of soft tissues which can allow discrimination of, for example, adipose from muscle from connective tissue from lung parenchyma. X-ray absorptivity is inherently quantifiable as Hounsfield units (HU), which permits advanced analytics, as well as ready comparisons across scanners and institutions. A limitation for the use of CT is that X-rays can also be absorbed by DNA, leading to double strand breaks. Even though modern scanners are built to mitigate this by reducing the amount of radiation needed to generate a useable image, the use of ionizing radiation in a potentially otherwise healthy population is a well-recognized limitation. The two major uses of X-rays in primary screening are in mammography for early detection of breast cancers, and low-dose CT (LDCT) for early

detection of lung cancers, both of which will be discussed in the subsequent section.

### MRI

MRI is advantageous in that it also has excellent tissue penetration and delivers superior soft tissue (and tumor) contrast without ionizing radiation. However, MRI generally requires long acquisition times and hence leads to motion artifacts, especially in abdominal tissues. This is being mitigated with advanced compressed sensing techniques, including MR fingerprinting, which have the potential to render MR images in seconds, compared with the current minutes. A mixed blessing of MRI is also the wide variety of pulse sequences available that can generate contrast based on a multitude of physicochemical magnetic relaxation mechanisms. While this makes MRI an extremely powerful technique, it is difficult to achieve uniformity of acquisition parameters across multiple institutions for a qualifying study. These limitations are considered challenges by the MR research community, and meeting these challenges will be discussed in the final section. In current practice, multiparametric MRI (mpMRI) has found a niche in the active surveillance (AS) of prostate cancers to discriminate indolent from progressing disease, and this will be discussed in a subsequent section.

### Nuclear

Nuclear imaging with single photon (SPECT) or positron (PET) tracers is not commonly associated with screening in an asymptomatic population. However, this is gaining traction in difficult to detect cancers, such as pancreatic, where the  $\alpha v \beta 6$  integrin has been shown to be an effective marker, with a highly specific 18-F positron-emitting tracers, and this will also be discussed in a subsequent section.

## Specific Use Cases

Specific use cases of cancer sites, although not meant to be exhaustive reviews were chosen to illustrate the breadth of quantitative imaging approaches in cancer screening and early detection. In all cases of cancer sites, we identify where imaging currently and potentially fits within early detection paradigms, and whenever possible, how imaging data can be combined with serum biomarkers for improved sensitivity (if primary screen) and specificity (if a secondary screen).

### Cutaneous neoplasms (optical)

The most obvious and impactful early detection setting that results in reduced cancer burden and mortality is optical detection and diagnosis of potentially lethal skin cancers, such as metastatic melanoma and squamous cell carcinomas. The accepted paradigm for assessing the risk of a dysplastic nevus is a validated quantitative scoring scheme known as “ABCD” for Asymmetry, Border irregularity, Color variegation, and Diameter >6 mm (19). Although this is a subjective scoring system, it is nonetheless effective and can be considered “quantitative imaging” (20). Automated or semiautomated quantification of the ABCD scoring system is an active area of research (21). In particular, because of the ready availability of annotated images, this is a particularly ripe area for the application of DL for classification. Deeply learned algorithms, such as CNNs do not rely on dermatologist-defined ABCD scoring criteria, but instead identify and develop informative features on their own.

In 2016, a large study trained a CNN on 129,450 annotated dermatologic images representing 2,032 different skin conditions,

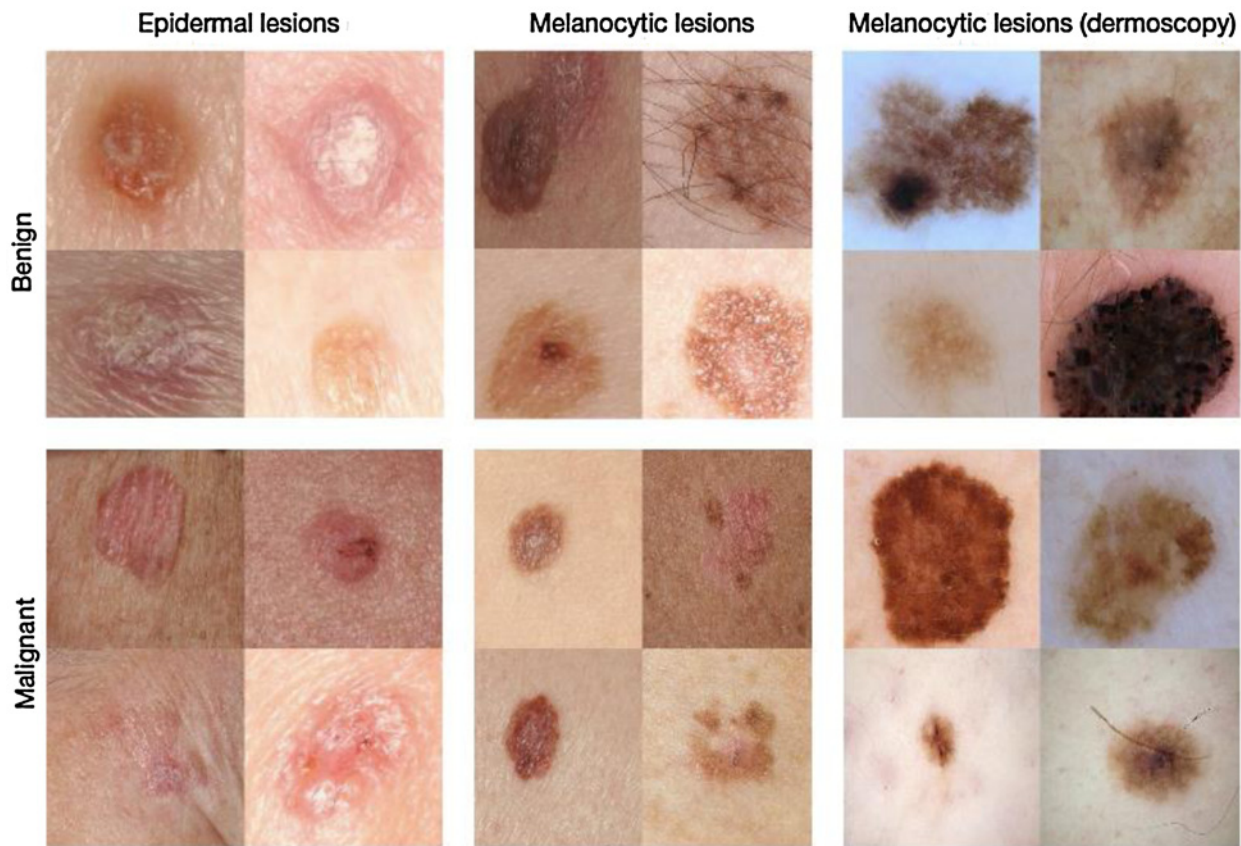
from benign rashes to malignant melanoma (22). Examples of the classification schema and the training images are provided in Fig. 2. In this study, the AI algorithm was then tested against 21 board-certified dermatologists in two clinically relevant binary classification tasks: keratinocytic carcinomas versus benign seborrheic keratoses, and malignant melanoma versus benign nevi. In both tasks, the AI algorithm achieved a performance level indistinguishable from the experts. More recently, deeper networks with further training have been shown to now outperforming dermatologists in classification accuracy (23). This paradigm of using AI to aid in differential diagnosis of skin lesions has been subjected to two large meta-analyses, including a Cochrane Review (24, 25). These concluded that AI approaches are highly sensitive for identification of invasive melanoma in select populations. In cutaneous lesions, the penalty for false positives is relatively low in that it triggers a simple excisional biopsy that is analyzed histologically. However, the biopsied sample is still limited by classification accuracy by micropathologists. Even here, DL algorithms are being shown to improve classification accuracy by dermatopathologists (26).

If these enterprises develop to the point of approval and acceptance, it has been imagined that AI applications, perhaps even on a smartphone, will be routinely used by patients, dermatologists, and dermatopathologists to routinely identify skin lesions before they have progressed to an incurable stage.

### Lung cancer (CT)

For a number of reasons, radiomics has been developed and applied in the diagnosis, prediction, and prognosis of lung cancer far more than any other disease. These reasons include the high disease incidence and subsequent large number of cases available, the high CT imaging contrast that naturally exists between lung nodules or tumors and the surrounding parenchyma, the ubiquity of CT scans in the workup and management of lung cancers, and the biomedical importance of the problem. In an early detection paradigm, lung nodules are detected using non-contrast-enhanced LDCT screening of high-risk individuals (27). However, because screening programs have not been embraced by high-risk smokers or former smokers, indeterminate pulmonary nodules are most commonly detected incidentally during a CT scans for other conditions, such as heart calcium scoring, trauma, etc. (28, 29). Screen-detected nodules are managed by Lung CT Screening Reporting & Data System (Lung-RADS) scoring (30), or increasingly using the Brock University cancer prediction model (31). Incidental nodules are generally managed using the Fleischner Society guidelines for the management of pulmonary nodules (32). All of these aforementioned schemes include CT-measured nodule size as an integral component. Size is one of the main indicators, in addition to growth rate over time, to determine the nature of a pulmonary nodule (33).

In the screening setting, the publicly available data and images from the National Lung Screening Trial (NLST; ref. 34) has enabled dozens of notable radiomic studies. In one of the first well-powered studies, a conventional radiomic score was developed using machine learning based on 10 features (size, shape, and texture) that were extracted from segmented baseline indeterminate (4–12 mm) pulmonary nodules (35). In independent testing, the radiomics score outperformed volume, and had virtually the same overall accuracy of the Brock model in predicting subsequent development of cancers within 1 or 2 years after the baseline screen. Notably, the radiomics model outperformed the Brock score with accuracies >0.93 for the extreme cases of very high and very low radiomics



**Figure 2.**

Malignant and benign example images from two disease classes. These test images highlight the difficulty of malignant versus benign discernment for the three medically critical classification tasks we consider: epidermal lesions, melanocytic lesions, and melanocytic lesions visualized with a dermoscope. Example images for **Fig. 2** are reprinted with permission from the Edinburgh Dermofit Library (available from: <https://licensing.eri.ed.ac.uk/i/software/dermofit-image-library.html>).

scores, which represented approximately 60% of the screening subjects.

More recently, a powerful study from Google Health developed an end-to-end detection and diagnosis system using a deep learned Convolutional Neural Net (36). This model was developed based on 42,290 LDCT scans from 14,851 patients curated from the NLST. Of these, 578 developed biopsy confirmed lung cancer within 1 year of follow up. Patients were randomized into training (70%), tuning (15%), and testing (15%) sets. A powerful component of machine learned models is that the cutoffs can be varied to generate different “lung malignancy scores” (LUMAS), for comparison with Lung-RADS scores from 1 to 2 to 4X. In all cases, the AI model outperformed scoring by experienced radiologists using Lung-RADS scoring schema. Moving forward, it will be important to test this predictive model in a real-world setting. Furthermore, additional datasets from other screening studies [MILD (37), LUSI (38), NELSON (39)] should become available furthering the ability to train, test, validate and improve the power of these and other models.

An important consideration of radiomics, as described above, is that the prediction is based on images at a single time-point, which may or may not contain additional clinical data. Going forward, it will be important to incorporate longitudinal image data, or delta radiomics, much as oncologists do in their decision processes. The temporal changes in radiomic features has great potential to inform

changes in the nodule biology that have important diagnostic and pathologic consequences (40, 41). It will also be important to incorporate quantitative biochemical information from serum or sputum in these models, as they clearly contain important orthogonal information (42, 43).

Despite the benefits of lung cancer earlier detection, LDCT screening identifies large numbers of false positives and indeterminate pulmonary nodules, of which only a fraction develops into cancer and LDCT screening detects indolent neoplasms that may not otherwise cause clinical symptoms or death (i.e., overdiagnosis). Because of these limitations, investigators in the EDRN have been developing image-based biomarkers along with genomic, molecular, and cellular biomarkers to improve diagnosis of nodules and reduce false positives and overtreatment. The Lung Cancer Collaborative Group brings together investigators from EDRN Clinical Validation Centers (CVC) and Biomarker Development Laboratories (BDL) with the primary focus is to develop and validate biomarkers and imaging methods to detect lung cancer among smokers with indeterminate nodules that are detected by LDCT.

**Breast cancer (mammography, mpMRI)**

Breast cancer screening by mammography has a long history. In screening for this disease, it is critical to have as high an accuracy as possible as the consequences of false positives leading to unnecessary

Downloaded from <http://aacrjournals.org/cebp/article-pdf/29/12/2556/1948353/2556.pdf> by guest on 28 August 2022

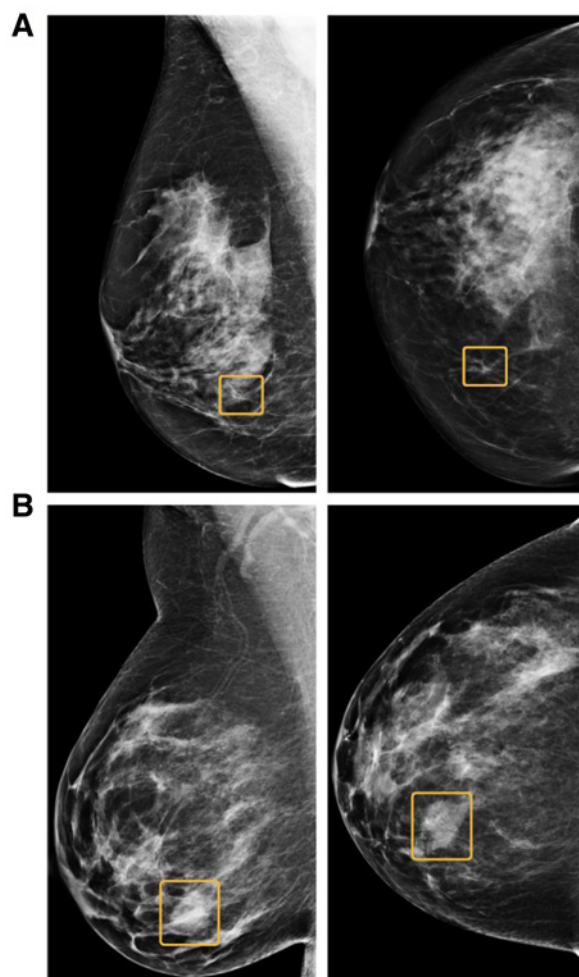
interventions, and false negatives leading to cancer progression, can be disastrous (44, 45). Because of its long history, there are vast repositories of outcomes-annotated mammographic images. Consequently, the use of these images to develop and deploy computer-aided diagnostic (CAD) workstations to aid in the interpretation of mammographic images has a similarly long history. CAD systems have been shown to work very well in controlled research settings, but they generally fell short in “real-world” clinical settings. Despite widespread use of CAD in clinical practice (46), a large multi-institutional study in 2015 of over 600,000 mammograms concluded that “CAD does not improve diagnostic accuracy of mammography” (47). Although breast radiologists are willing to embrace technology, there is skepticism by some regarding the value of newer AI-driven algorithms in routine clinical practice (48).

This may soon change, however, with the publication of a new international study by Google Health (49). This study trained and tested a CNN on mammograms from 28,953 women in the United States and the United Kingdom who had at least 1-year of follow up to classify mammograms as showing cancer (biopsy-proven) versus noncancer (either biopsy proven or no lesion). In this study, the AI algorithm outperformed all of the radiologists who had originally read the mammograms, as well as 6 expert radiologists who prospectively over-read 500 cases. The AI algorithm had significantly higher sensitivities and specificities over human readers to predict emergence of cancer within 3 years (United Kingdom) or 2 years (United States). Nonetheless, the AI system did contain some false negatives that were identified as positive by radiologists, and vice versa. Examples of these are given in Fig. 3A and B. This complementarity suggests that combining AI with radiology reads would provide the highest accuracy and likely lead to more ready acceptance.

Similar to the efforts related to lung cancer, EDRN investigators have been combining image-based and circulating biomarkers to improve diagnostic discrimination in breast cancer screen and to discriminate between indolent and aggressive subtypes. The EDRN Breast and Gynecologic Cancers Collaborative Group brings together CVCs and BDLs to conduct research to improve the performance of screening mammography, to distinguish benign from malignant breast lesions, and improve early detection of different molecular subtypes of breast cancer.

### Pancreatic cancer (CT, MRI, PET/CT)

Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal of human cancers and may become the most lethal by 2022 (50). Whereas the overall 5-year survival rate is 6% to 7% (51), the survival rate is virtually zero for disease that is unresectable (52–56). Because of this, pancreatic cancer has a most compelling need for an early detection portfolio, as this cancer is commonly diagnosed at an advanced stage, when it is lethal. As such, this is an area of intense investigation by combining serum and image biomarkers. In contrast to other cancers, there are no clearly defined schema for detecting pancreatic lesions in an asymptomatic population. Patients at high risk for development of PDAC include those with family history, symptomatic pancreatitis, new-onset diabetes (NOD), or asymptomatic mucinous cystic lesions (cystic neoplasms and intra papillary mucinous cystic neoplasm; IPMN), and thus are candidates for active surveillance with endoscopic US (EUS) or MR imaging (57). The other precursor lesion, pancreatic intraepithelial neoplasia, or PanIN, is rarely, if at all, incidentally detected in asymptomatic individuals. Cystic lesions are widely detected incidentally in approximately 1.2% of all abdominal CTs, or approximately 150,000 annually (58). Although



**Figure 3.**

A schematic illustration of the taxonomy and example test set images. **A**, A sample cancer case that was missed by all six readers in the U.S. reader study, but correctly identified by the AI system. The malignancy, outlined in yellow, is a small, irregular mass with associated microcalcifications in the lower inner right breast. **B**, A sample cancer case that was caught by all six readers in the U.S. reader study, but missed by the AI system. The malignancy is a dense mass in the lower inner right breast. Left, mediolateral oblique view; right, craniocaudal view. **Figure 3** is with permission from McKinney and colleagues (available from: <https://www.nature.com/articles/s41586-019-1799-6?proof=true>).

IPMNs are precursor lesions in only approximately 15% of all PDAC, they are illustrative of the early detection paradigm, as they are difficult to manage, as <1% of all cystic lesions are destined to progress to PDAC (59). Current management of IPMNs involves surgery, which is associated with both operative mortality and multiple morbidities, as well as overtreatment, as 50% to 70% of resected cystic lesions are nonmalignant (60). The decision to resect is often based on radiographic identification of “worrisome” or “high-risk” features. Patients not electing surgery have an option of active surveillance with biannual EUS or MRI (61). Historically, there is poor agreement between preoperative diagnosis and pathologic examination, identifying a great need for improved quantitative markers of malignancy risk (62–67).

The current paradigm for early detection of PDAC is primarily based on biomarkers present in serum, and this is an active area of

investigation (68, 69). The matrix glycoprotein thrombospondin-2 regulates remodeling of the extracellular matrix (ECM), and hence its presence in serum is indicative of ECM breakdown and resynthesis (70). Recently, it has been proposed as a biomarker to identify IPMNs that are at high risk for progression to PDAC (71), and the power of this prediction was enhanced by combining with CA 19-9 levels and qualitative radiographic features (72). In addition, circulating miRNAs have also been shown to identify IPMNs that are destined to progress (73). In a follow-up study, the “miRNA genomic classifier” (MGC) and quantitatively scored radiologic features were extracted from a cohort of 38 surgically resected, pathologically confirmed IPMN cases. The AUROC of the MGC to predict malignancy was 0.83, and the high-risk radiologic features yielded an AUROC of 0.84. Fourteen textural and nontextural radiomic features differentiated malignant from benign IPMNs ( $P < 0.05$ ) and collectively had an AUROC of 0.77. Combining these textural radiomic features with the MGC yielded an AUROC of 0.92. Sensitivity (83%), specificity (89%), positive predictive value (88%), and negative predictive value (85%) were superior to conventional clinical pathologic models (74). Significant findings of this article were that radiomics could identify true negatives from otherwise false positives (Fig. 4A) and that there was one case that was true positive that was correctly classified by radiomics but was a false negative by conventional qualitative imaging (Fig. 4B).

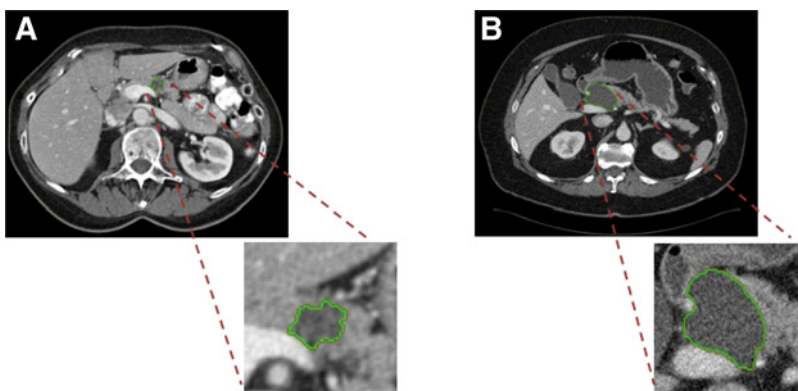
Currently, there are multiple investigations to apply AI to abdominal images with the expectation that DL can identify precursor lesions before they are evident to radiologists (75). As with much of radiomics, this effort suffers from availability of multi-institutional, well-curated cohorts of past CT scans from current PDAC patients. Hence, most efforts in AI analyses of PDAC have been focused on discriminating invasive from benign neoplasms. For example, Koay and colleagues (MD Anderson Cancer Center, Houston, TX) have shown that quantitative analysis of the tumor-stromal interface was a significant prognostic indicator in 3 different cohorts ( $N = 303$  total) of patients with PDAC treated with surgery or chemotherapy (76). Chu and colleagues (Felix project at Johns Hopkins, Baltimore, MD) have related their initial investigations using deep CNNs to identify normal pancreas (and other visceral tissues) as well as classifying PDAC as a case study (77). They trained the network on 575 control subjects and were able to achieve excellent segmentations of most major organs with accuracies  $>88\%$  (compared with manual segmentations). This is an important first step, as it can be expected that precursor lesions will deviate from the norm. The network was also trained to recognize PDAC as a deviation from norm, and this involved a cohort of 750 patients with PDAC. Even though the acquisition parameters of this

training set were homogeneous, the model performed well only on larger lesions. Hence, its eventual application to the identification and characterization of small precursor lesions remains to be demonstrated. Although this is the right approach, it will require significantly larger and more heterogeneous training and testing sets curated from prior scans from confirmed patients with PDAC.

Because asymptomatic pancreatic cystic lesions are primarily identified incidentally, more structured risk assessment tools are critically needed. A typical screening paradigm is unworkable due to the relatively low prevalence, even though additional risk factors are being identified (57). Because of low prevalence, serum biomarkers have also been elusive, and this is an area of active investigation within the EDRN network. CA 19-9 is the most common PDAC-associated antigen, yet it does not become measurable until disease has progressed. Combining CA 19-9 with autoantibodies, circulating antigens, miRNA, or enzymes such as cathepsins, however, may have promise to improve its sensitivity to incipient disease (68, 78–80). It is reasonable to expect that, in the foreseeable future, serum biomarkers to identify patients “at-risk” of incipient pancreatic hyperplasia to create a cohort with a tractable prevalence, which then could be used to trigger assessment by secondary imaging/radiomics-based screens.

When radiomic analyses are incorporated in differential diagnoses, improvements in sensitivity and specificity have been shown in determining risk of malignancy. EDRN investigators have been actively conducting research utilizing imaging and molecular biomarkers to develop a rational, evidence-based strategy to detect pancreatic cancer at an early and resectable stage. The EDRN GI Collaborative Group conducts research on colorectal, pancreatic, and esophageal cancers and there are two CVCs and one BDL focused on pancreatic cancer. The main goals of this Collaborative Group with respect to pancreatic cancer to develop and validate biomarkers and imaging methods to detect pancreatic cancer in high-risk groups and to develop and validate biomarkers and imaging methods to determine which pancreatic cysts are cancerous. In recognition of the importance of images to this enterprise, the PDAC working group is engaged in the construction of an image repository of early-stage PDAC and IPMNs.

Nuclear medicine approaches, such as PET, are usually not included within an early detection pipeline due to high costs and choice of tracers. However, this may be changing with pancreatic cancers, due to the severity of the problem and the need to identify early lesions with high risk of malignancy. Through many studies, the  $\alpha\beta6$  integrin has been shown to be highly associated with incipient pancreatic cancers and has been shown to be a driver of PDAC



**Figure 4.**

Semiautomated segmentation of two IPMN patient CT scans at selected central slice. **A**, Axial venous phase images through the abdomen demonstrate a cystic mass in the pancreatic head/neck measuring up to 3.5 cm. This lesion contains a nonenhancing soft tissue mural nodule (arrow). **B**, Axial venous phase images through the abdomen demonstrate an ovoid, homogeneous appearing cystic mass measuring up to 4.8 cm in greatest dimension. No internal enhancing soft tissue nodules were seen within the lesion. Images from Fig. 4 are with permission from Permut and colleagues (available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5349874/>).

phenotypes (81). As such, specific PET tracer labeled ligands for integrin  $\alpha v \beta 6$  have been developed (82, 83). These tracers have been used in humans, and are proposed to be used in a screening paradigm (84), which would be the one of the first instances of such use for PET imaging.

#### Prostate cancer (mpMRI and tumor habitats)

Prostate cancer is one of the most commonly diagnosed cancers and accounts for the second largest number of cancer-related deaths among men (85). Most patients with prostate cancer are first identified by rising serum-based prostate-specific antigen (PSA) coupled with a digital rectal exam (86, 87). In contrast to other use cases described above, the application of imaging AI to Prostate Cancer occurs after cancer is detected, usually with elevated or rising PSA, even though this risk marker is limited by low sensitivity and specificity (87, 88). Unlike other cancers, prostate cancer is often detected early and can often be indolent, posing no risk during a subject's life time. However, if prostate cancer is high grade, with Gleason scoring of 4+3 or above, it has the potential to be life threatening. Therefore, discriminating indolent from aggressive disease is of utmost clinical importance.

Increasingly, imaging occurs in a setting of AS, which has emerged as an alternative for patients with localized, low-grade (Gleason <7) disease, where treatment options such as radiotherapy or prostatectomy along with their accompanying morbidity that could be postponed, by as much as 5 years without significant change in outcome (89–91). In AS schema, men with low-grade disease are regularly (annual or biannual) imaged with multiparametric mpMRI, consisting of a diffusion-weighted sequences, T2 images for morphology, and a dynamic contrast series following a bolus injection of a contrast agent. Coincident with these imaging sessions, men are routinely subjected to transrectal US (TRUS)-guided biopsy, generally consisting of 12 to 14 random cores along with one to two targeted biopsies of suspicious regions identified on the mpMRI. Current practice is to analyze mpMR images of the prostate using the Pi-

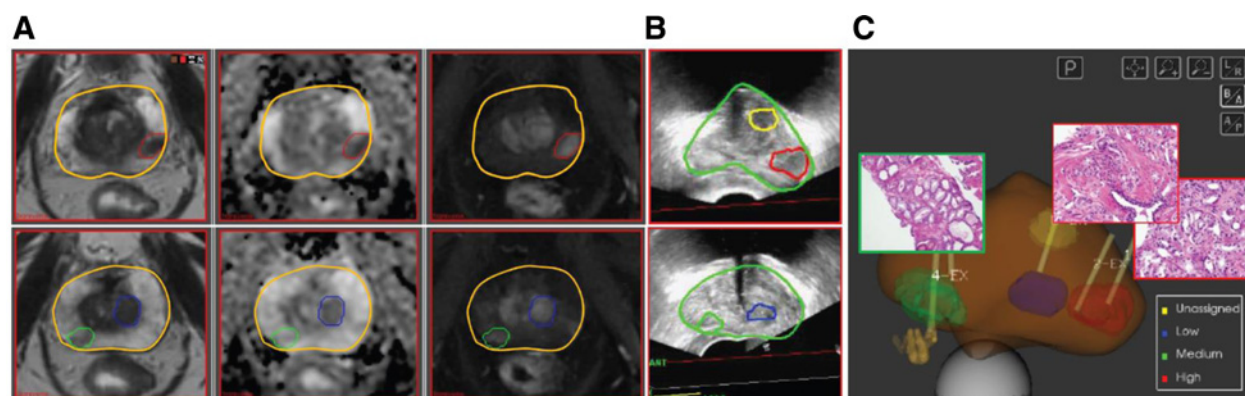
RADS scoring system (92–94). However, Pi-RADS suffers from poor inter-reader concordance (95). Hence, there is an interest in developing an automated (AI) strategy that offers greater reliability than is achieved under the current paradigm (96).

Advances in imaging technologies have made it possible to obtain high-resolution functional imaging data that are capable of reflecting tumor cellularity, dynamics of contrast uptake, anatomical location and visual characteristics of the lesion in the entire gland in 3-dimensions. In recent years, investigators have been developing methods to combine image data from multiple image data sets to identify spatially distinct regions with similar physiologies, or “habitats” (4). In early prostate cancer, these habitat imaging approaches have shown promise in identification of “at-risk” volumes within the prostate that can be used to guide TRUS biopsy locations, as shown in Fig. 5 (97–99), or to improve risk stratification compared to Pi-RADS (100). The use of radiomics and habitat imaging in the characterization of early prostate has recently been reviewed (101, 102). The long-term goal of radiomics and habitat imaging in early prostate cancers is to gain enough confidence that the imaging method to detect progressive disease, so that annual core biopsies in an AS schema can be eliminated.

In the EDRN, investigators have been conducting research to improve the early detection of prostate cancer, for prostate cancer upgrading, and fusion MRI to improve the accuracy of prostate needle biopsies. The EDRN Prostate and Other Urological Cancers Collaborative Group is developing noninvasive tests to discern indolent cancers from aggressive cancers and to determine whether MRI prostate imaging and biomarkers can improve the prediction of cancer extent and aggressiveness to determine suitability for active surveillance or treatment.

#### Liver cancer (US elastography)

Liver fibrosis, cirrhosis, nonalcoholic steatohepatitis (NASH), and nonalcoholic fatty liver disease (NAFL) are increasingly common pathologic conditions, each with their own complications (103, 104).



**Figure 5.**

Delineation of biopsy targets on mpMRI and fusion of targets on 3D TRUS. **A**, Screenshots from ProFuse software (Eigen) for fusion of mpMRI delineated prostate ROIs for targeted biopsy. Two axial slices, going from base (top) to apex (bottom), are displayed. The prostate volume is outlined (yellow contour); (left) T2-weighted MRI; (center) apparent diffusion coefficient (ADC) derived from diffusion weighted imaging (DWI); and (right) early enhancing image from dynamic contrast-enhanced (DCE-)MRI. The volumes in red, green, and blue are assigned high, medium, and low probability for cancer. **B**, A screenshot from Artemis (Eigen), displaying the 3D TRUS views corresponding to the axial slices in **(A)** after nonrigid fusion of the prostate boundaries on MRI and US. The targets are transferred from mpMRI to real-time US biopsy system. **C**, Schematic representation of the prostate and target volumes. (Note that the display contains an ROI in yellow is with unassigned probability). Yellow lines indicate biopsy needle tracks (1 needle in green, 2 in red, and 1 in blue); the corresponding hematoxylin and eosin slides at 20 $\times$  magnification from green target (left; Gleason score 6) and red targets (right; Gleason score 7). Images for Fig. 5 are with permission from Stoyanova and colleagues (available from: <https://www.ncbi.nlm.nih.gov/pubmed/27438142>).

Notably, all of these are predisposing conditions for development of hepatocellular carcinoma (HCC; ref. 105). Historically, HCC has been associated with hepatitis (Hep) virus B or C infections leading to fibrotic liver disease. Thus, Hep B and C status are powerful risk factors to enter into a screening program. However, with the uptake of vaccinations, and increased prevalence of obesity, fatty liver disease is increasingly being recognized as a relevant risk factor as well (106). Obesity-associated NAFL is generally benign, yet a subset can develop into NASH, which is at high risk for developing cirrhosis (107). The gold standard to diagnose NASH versus NAFL is liver biopsy, but this is associated with morbidity and even mortality. Clinical predictors and serum biomarkers have been developed, yet they have not been validated in multi-institutional case-control studies (107). Hence, noninvasive measures are critically needed.

Because of the prevalence of predisposing conditions, increasing numbers of patients are routinely having their livers examined by acoustic force (ARFI) or shear-wave elastography (SWE), which are widely available (108). While these approaches are not generally associated with “Radiomics” per se, there is increasing analysis of US with DL (109). Notably, the risk of developing HCC increases with increasing hepatic stiffness values, reported quantitatively in kilopascals (kPa). In a well-powered and highly cited study of 866 Japanese patients with Hep C infection, Masuzaki and colleagues observed that the incidence of HCC within 3 years was 38.5% among those with baseline liver stiffness values >25 kPa, compared with 0.4% among subjects with values ≤10 kPa (110). The strengths and limitations of elastography for diagnosis of fibrosis have been extensively reviewed in a consensus conference, with the conclusion that it can eliminate the need for liver biopsy in most patients (111).

If more detailed examinations are needed prior to interventions, it is also possible to perform elastography using MRI, wherein the body is mechanically stimulated with low-frequency vibrations, and the resulting movement of underlying tissues is measured with the MRI signal (112). Although MR elastography is less available compared with US elastography, it does have the advantage of providing a larger field of view to encompass the entire abdomen during a single imaging session. **Figure 6** shows examples of MR elastographic images from patients at low, intermediate, and high risk of cirrhosis and hence HCC.

Although the EDRN does not have CVCs or BDLs focused on liver cancer, the network has begun prospectively collecting samples as part

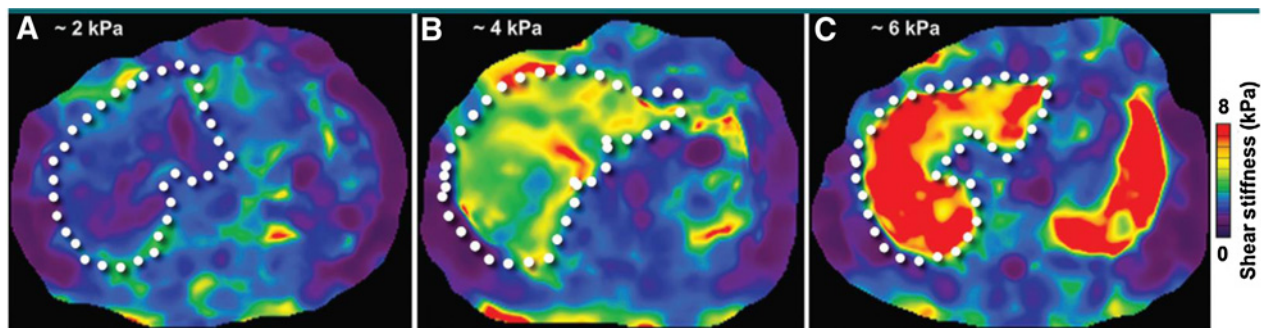
of the Hepatocellular Carcinoma Early Detection Strategy Study (HEDS). The goal of these efforts is to define better risk stratification, which could help identify additional high-risk patients that need triaged to surveillance programs and that identifying and improving the surveillance of patients at risk may help reduce mortality due to liver cancer.

## Conclusions, Challenges, and Future Directions

Images are data. As described above, quantitative analyses of these data can yield highly predictive models based on a parsimonious set of informative imaging features. Conventional radiomics generates machine learned models using with supervised identification of the ROIs, as well as extraction using user-defined features. These have resulted in predictive models with very high accuracies. In nearly all cases, accuracy is improved when combining the radiomics model with orthogonal clinically derived data, such as serum markers, demographics, histopathology, genomics, etc. More recently, enabled by the availability of large publicly available and annotated datasets, images are being analyzed using DL approaches. In DL, the computer can analyze the entire image or automatically detect the ROI, or an operator can focus the computer's attention on the general region via, for example, a bounding box. In neural nets, the computer then identifies image features that are most informative to generate a classification. More detailed reviews on DL and AI in imaging are available elsewhere (3–6).

In either case, conventional or deeply learned radiomic training sets require large amounts of well-curated data. Curation requires identification of clean outcomes which are then used in the classifier tasks. This brings up two challenges, *viz.* very large datasets are not readily available, and outcomes data (e.g., cancer vs. noncancer; low grade vs. high grade) generally have to be manually defined and curated in a structured format. In cases where very large curated datasets have been generated, for example, the NLST, the image acquisition protocols that were used then do not represent current technology and clinical practices. Correcting for this presents its own challenges.

A future solution to solve most of these issues is the adoption of a vast distributed learning network, wherein institutions curate and retain their own data, and algorithms are shared to improve training,



**Figure 6.**

MR elastograms in three patients. **A**, MR elastogram in a 63-year-old woman with a history of autoimmune hepatitis shows that the liver (outlined) has no evidence of hepatic fibrosis, with a normal stiffness value of approximately 2 kPa. **B**, MR elastogram in a 52-year-old woman with chronic cholestatic hepatitis demonstrates increased hepatic stiffness with approximately twice the normal value at 4 kPa, indicating the presence of significant hepatic fibrosis. **C**, MR elastogram in a 46-year-old man with chronic hepatitis C infection demonstrates markedly increased hepatic stiffness, averaging over 6 kPa, as is consistent with the presence of advanced hepatic fibrosis (cirrhosis). Images from **Fig. 6** are with permission from Barr and colleagues (available from: <https://pubs.rsna.org/doi/full/10.1148/radiol.2015150619>).

testing, and validation (11). A simplified extension of this would be for journals to require deposition of code (whether compiled or raw) in an accessible repository so that promising models could be tested and replicated at different sites. Notably, two of the highest profile recent papers used DL to predict lung nodule status from LDCT scans (36), and detection of breast cancers from mammography (49). In both cases the authors refrained from sharing code, with the rationale that: “code used for training the models has a large number of dependencies on internal tooling, infrastructure and hardware, and its release is therefore not feasible.” This is unacceptable, particularly if the training data were generated by public funds.

## References

1. Srivastava S, Koay EJ, Borowsky AD, De Marzo AM, Ghosh S, Wagner PD, et al. Cancer overdiagnosis: a biological challenge and clinical dilemma. *Nat Rev Cancer* 2019;19:349–58.
2. Grizzle WE, Srivastava S, Manne U. Translational pathology of neoplasia. *Cancer Biomark* 2010;9:7–20.
3. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2015;278:563–77.
4. Napel S, Mu W, Jardim-Perassi BV, Aerts H, Gillies RJ. Quantitative imaging of cancer in the postgenomic era: Radio(geno)mics, deep learning, and habitats. *Cancer* 2018;124:4633–49.
5. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J Clin* 2019;69:127–57.
6. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer* 2018;18:500–10.
7. Fave X, Zhang L, Yang J, Mackin D, Balter P, Gomez DR, et al. Using pretreatment radiomics and delta-radiomics features to predict non-small cell lung cancer patient outcomes. *Int J Radiat Oncol Biol Phys* 2017;98:249.
8. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning: a primer for the epidemiologist. *Am J Epidemiol* 2019;188:2222–39.
9. Shen D, Wu G, Suk H-I. Deep learning in medical image analysis. *Annu Rev Biomed Eng* 2017;19:221–48.
10. Xue Y, Chen S, Qin J, Liu Y, Huang B, Chen H. Application of deep learning in automated analysis of molecular images in cancer: a survey. *Contrast Media Mol Imaging* 2017;2017:9512370.
11. Chang K, Balachandrar N, Lam C, Yi D, Brown J, Beers A, et al. Distributed deep learning networks among institutions for medical imaging. *J Am Med Inform Assoc* 2018;25:945–54.
12. O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* 2017;14:169–86.
13. Liu Y, Kim J, Qu F, Liu S, Wang H, Balagurunathan Y, et al. CT features associated with epidermal growth factor receptor mutation status in patients with lung adenocarcinoma. *Radiology* 2016;280:271–80.
14. Grossmann P, Stringfield O, El-Hachem N, Bui MM, Rios Velazquez E, Parmar C, et al. Defining the biological basis of radiomic phenotypes in lung cancer. *Elife* 2017;6:e23421.
15. Mu W, Tunalı I, Gray JE, Qi J, Schabath MB, Gillies RJ. Radiomics of (18)F-FDG PET/CT images predicts clinical benefit of advanced NSCLC patients to checkpoint blockade immunotherapy. *Eur J Nucl Med Mol Imaging* 2020;47:1168–82.
16. Cheng K, Cheng Z. Near infrared receptor-targeted nanoprobes for early diagnosis of cancers. *Curr Med Chem* 2012;19:4767–85.
17. Tang Q, Wang J, Frank A, Lin J, Li Z, Chen CW, et al. Depth-resolved imaging of colon tumor using optical coherence tomography and fluorescence laminar optical tomography. *Biomed Opt Express* 2016;7:5218–32.
18. Tanaka T, Kamata M, Fukaya S, Hayashi K, Fukuyasu A, Ishikawa T, et al. Usefulness of real-time elastography for diagnosing lymph node metastasis of skin cancer: does elastography potentially eliminate the need for sentinel lymph node biopsy in squamous cell carcinoma? *J Eur Acad Dermatol Venereol* 2020;34:754–61.
19. Ahnslide I, Bjellerup M, Nilsson F, Nielsen K. Validity of ABCD rule of dermoscopy in clinical practice. *Acta Derm Venereol* 2016;96:367–72.
20. Glazer AM, Farberg AS, Winkelmann RR, Tucker N, Rigel DS. Quantitative ABCD parameters measured by a multispectral digital skin lesion analysis device for evaluation of suspicious pigmented skin lesions strongly correlate with clinical ABCD observations. *Dermatol Online J* 2017;23:1–4.
21. Thanh DNH, Prasath VBS, Hieu LM, Hien NN. Melanoma skin cancer detection method based on adaptive principal curvature, colour normalisation and feature extraction with the ABCD rule. *J Digit Imaging* 2020;33:574–85.
22. Esteve A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115–8.
23. Brinker TJ, Hekler A, Enk AH, Berking C, Haferkamp S, Hauschild A, et al. Deep neural networks are superior to dermatologists in melanoma image classification. *Eur J Cancer* 2019;119:11–7.
24. Ferrante di Ruffano L, Takwoingi Y, Dinnes J, Chuchu N, Bayliss SE, Davenport C, et al. Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults. *Cochrane Database Syst Rev* 2018;12:CD013186.
25. Marchetti MA, Liopyris K, Dusza SW, Codella NCF, Gutman DA, Helba B, et al. Computer algorithms show potential for improving dermatologists' accuracy to diagnose cutaneous melanoma; results of ISIC 2017. *J Am Acad Dermatol* 2020;82:622–7.
26. Hekler A, Utikal JS, Enk AH, Solass W, Schmitt M, Klode J, et al. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. *Eur J Cancer* 2019;118:91–6.
27. U.S. Preventive Services Task Force. Final update summary: lung cancer. Rockville (MD): USPSTF; 2015. Available from: <https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/lung-cancer-screening>.
28. Scholtz JE, Lu MT, Hedgire S, Meyersohn NM, Oliveira GR, Prabhakar AM, et al. Incidental pulmonary nodules in emergent coronary CT angiography for suspected acute coronary syndrome: impact of revised 2017 Fleischner Society Guidelines. *J Cardiovasc Comput Tomogr* 2018;12:28–33.
29. Hammerschlag G, Cao J, Gumm K, Irving L, Steinfort D. Prevalence of incidental pulmonary nodules on computed tomography of the thorax in trauma patients. *Intern Med J* 2015;45:630–3.
30. American College of Radiology. Lung CT screening reporting & data system (Lung-RADS). Reston (VA): ACR. Available from: <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads>.
31. McWilliams A, Tammemagi MC, Mayo JR, Roberts H, Liu G, Soghrati K, et al. Probability of cancer in pulmonary nodules detected on first screening CT. *N Engl J Med* 2013;369:910–9.
32. MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology* 2017;284:228–43.
33. Larici AR, Farchione A, Franchi P, Ciliberto M, Cicchetti G, Calandriello L, et al. Lung nodules: size still matters. *Eur Respir Rev* 2017;26:1–16.
34. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365:395–409.

35. Hawkins S, Wang H, Liu Y, Garcia A, Stringfield O, Krewer H, et al. Predicting malignant nodules from screening CT scans. *J Thorac Oncol* 2016;11:2120–8.
36. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 2019;25:1319.
37. Pastorino U, Silva M, Sestini S, Sabia F, Boeri M, Cantarutti A, et al. Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. *Ann Oncol* 2019;30:1162–9.
38. Becker N, Motsch E, Trotter A, Heussel CP, Dienemann H, Schnabel PA, et al. Lung cancer mortality reduction by LDCT screening—results from the randomized German LUSI trial. *Int J Cancer* 2020;146:1503–13.
39. De Koning H, Van der Aalst C, Ten Haaf K, Oudkerk M. Effects of volume CT lung cancer screening: mortality results of the NELSON randomised-controlled population based trial. *J Thorac Oncol* 2018;13:S185.
40. Alahmari SS, Cherezov D, Goldgof D, Hall L, Gillies RJ, Schabath MB. Delta radiomics improves pulmonary nodule malignancy prediction in lung cancer screening. *IEEE Access* 2018;6:77796–806.
41. Cherezov D, Hawkins SH, Goldgof DB, Hall LO, Liu Y, Li Q, et al. Delta radiomic features improve prediction for lung cancer incidence: a nested case-control analysis of the National Lung Screening Trial. *Cancer Med* 2018;7:6340–56.
42. Hulbert A, Jusue-Torres I, Stark A, Chen C, Rodgers K, Lee B, et al. Early detection of lung cancer using DNA promoter hypermethylation in plasma and sputum. *Clin Cancer Res* 2017;23:1998–2005.
43. Szepechinski A, Rudzinski P, Kupis W, Langfort R, Orlowski T, Chorostowska-Wynimko J. Plasma cell-free DNA levels and integrity in patients with chest radiological findings: NSCLC versus benign lung nodules. *Cancer Lett* 2016;374:202–7.
44. Wapnir IL, Dignam JJ, Fisher B, Mamounas EP, Anderson SJ, Julian TB, et al. Long-term outcomes of invasive ipsilateral breast tumor recurrences after lumpectomy in NSABP B-17 and B-24 randomized clinical trials for DCIS. *J Natl Cancer Inst* 2011;103:478–88.
45. Lari SA, Kuerer HM. Biological markers in DCIS and risk of breast recurrence: a systematic review. *J Cancer* 2011;2:232–61.
46. Keen JD, Keen JM, Keen JE. Utilization of computer-aided detection for digital screening mammography in the United States, 2008 to 2016. *J Am Coll Radiol* 2018;15:44–8.
47. Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL, et al. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med* 2015;175:1828–37.
48. Pisano ED. AI shows promise for breast cancer screening. *Nature* 2020;577:35–6.
49. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature* 2020;577:89–94.
50. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin* 2014;64:9–29.
51. American Cancer Society. Cancer facts and figures 2015. Atlanta (GA): American Cancer Society; 2015.
52. Ferrone CR, Brennan MF, Gonen M, Coit DG, Fong Y, Chung S, et al. Pancreatic adenocarcinoma: the actual 5-year survivors. *J Gastrointest Surg* 2008;12:701–6.
53. Howard TJ, Krug JE, Yu J, Zyromski NJ, Schmidt CM, Jacobson LE, et al. A margin-negative R0 resection accomplished with minimal postoperative complications is the surgeon's contribution to long-term survival in pancreatic cancer. *J Gastrointest Surg* 2006;10:1338–45.
54. Angst E, Kim-Fuchs C, Kuruvilla YC, Inderbitzin D, Montani M, Candinas D, et al. How to counter the problem of R1 resection in duodenopancreatectomy for pancreatic cancer? *J Gastrointest Surg* 2012;16:673.
55. Esposito I, Kleeff J, Bergmann F, Reiser C, Herpel E, Friess H, et al. Most pancreatic cancer resections are R1 resections. *Ann Surg Oncol* 2008;15:1651–60.
56. Campbell F, Smith RA, Whelan P, Sutton R, Raraty M, Neoptolemos JP, et al. Classification of R1 resections for pancreatic cancer: the prognostic relevance of tumour involvement within 1 mm of a resection margin. *Histopathology* 2009;55:277–83.
57. Singhi AD, Koay EJ, Chari ST, Maitra A. Early detection of pancreatic cancer: opportunities and challenges. *Gastroenterology* 2019;156:2024–40.
58. Farrell JJ. Prevalence, diagnosis and management of pancreatic cystic neoplasms: current status and future directions. *Gut and liver* 2015;9:571–89.
59. Pereira SP, Oldfield L, Ney A, Hart PA, Keane MG, Pandol SJ, et al. Early detection of pancreatic cancer. *Lancet Gastroenterol Hepatol* 2020;5:698–710.
60. Hines OJ, Reber HA. Pancreatic surgery. *Curr Opin Gastroenterol* 2008;24:603–11.
61. European Study Group on Cystic Tumours of the Pancreas. European evidence-based guidelines on pancreatic cystic neoplasms. *Gut* 2018;67:789–804.
62. Tanaka M, Fernandez-del Castillo C, Adsay V, Chari S, Falconi M, Jang JY, et al. International Consensus Guidelines 2012 for the management of IPMN and MCN of the pancreas. *Pancreatol* 2012;12:183–97.
63. Kim KW, Park SH, Pyo J, Yoon SH, Byun JH, Lee MG, et al. Imaging features to distinguish malignant and benign branch-duct type intraductal papillary mucinous neoplasms of the pancreas: a meta-analysis. *Ann Surg* 2014;259:72–81.
64. Roch AM, Ceppa EP, DeWitt JM, Al-Haddad MA, House MG, Nakeeb A, et al. International Consensus Guidelines parameters for the prediction of malignancy in intraductal papillary mucinous neoplasm are not properly weighted and are not cumulative. *HPB (Oxford)* 2014;16:929–35.
65. Sahara K, Mino-Kenudson M, Brugge W, Thayer SP, Ferrone CR, Sahani D, et al. Branch duct intraductal papillary mucinous neoplasms: does cyst size change the tip of the scale? A critical analysis of the revised international consensus guidelines in a large single-institutional series. *Ann Surg* 2013;258:466–75.
66. Fritz S, Klaus M, Bergmann F, Strobel O, Schneider L, Werner J, et al. Pancreatic main-duct involvement in branch-duct IPMNs: an underestimated risk. *Ann Surg* 2014;260:848–55.
67. Goh BK, Tan DM, Ho MM, Lim TK, Chung AY, Ooi LL. Utility of the Sendai Consensus Guidelines for branch-duct intraductal papillary mucinous neoplasms: a systematic review. *J Gastrointest Surg* 2014;18:1350–7.
68. Song J, Sokoll LJ, Pasay JJ, Rubin AL, Li H, Bach DM, et al. Identification of serum biomarker panels for the early detection of pancreatic cancer. *Cancer Epidemiol Biomarkers Prev* 2019;28:174–82.
69. Young MR, Wagner PD, Ghosh S, Rinaudo JA, Baker SG, Zaret KS, et al. Validation of biomarkers for early detection of pancreatic cancer: summary of the Alliance of Pancreatic Cancer Consortia for Biomarkers for Early Detection workshop. *Pancreas* 2018;47:135–41.
70. Calabro NE, Barrett A, Chamorro-Jorganes A, Tam S, Kristofik NJ, Xing H, et al. Thrombospondin-2 regulates extracellular matrix production, LOX levels, and cross-linking via downregulation of miR-29. *Matrix Biol* 2019;82:71–85.
71. Peng HY, Chang MC, Hu CM, Yang HI, Lee WH, Chang YT. Thrombospondin-2 is a highly specific diagnostic marker and is associated with prognosis in pancreatic cancer. *Ann Surg Oncol* 2019;26:807–14.
72. Simpson RE, Yip-Schneider MT, Wu H, Fan H, Liu Z, Korc M, et al. Circulating Thrombospondin-2 enhances prediction of malignant intraductal papillary mucinous neoplasm. *Am J Surg* 2019;217:425–8.
73. Permuth JB, Chen DT, Yoder SJ, Li J, Smith AT, Choi JW, et al. Linc-ing Circulating long non-coding RNAs to the diagnosis and malignant prediction of intraductal papillary mucinous neoplasms of the pancreas. *Sci Rep* 2017;7:10484.
74. Permuth JB, Choi J, Balarunathan Y, Kim J, Chen DT, Chen L, et al. Combining radiomic features with a miRNA classifier may improve prediction of malignant pathology for pancreatic intraductal papillary mucinous neoplasms. *Oncotarget* 2016;7:85785–97.
75. Pelaez-Luna M, Takahashi N, Fletcher JG, Chari ST. Resectability of presymptomatic pancreatic cancer and its relationship to onset of diabetes: a retrospective review of CT scans and fasting glucose values prior to diagnosis. *Am J Gastroenterol* 2007;102:2157–63.
76. Koay EJ, Lee Y, Cristini V, Lowengrub JS, Kang Y, Lucas FAS, et al. A visually apparent and quantifiable CT imaging feature identifies biophysical subtypes of pancreatic ductal adenocarcinoma. *Clin Cancer Res* 2018;24:5883–94.
77. Chu LC, Park S, Kawamoto S, Wang Y, Zhou Y, Shen W, et al. Application of deep learning to pancreatic cancer detection: lessons learned from our initial experience. *J Am Coll Radiol* 2019;16:1338–42.
78. Ye H, Wang H, Wang P, Song CH, Wang KJ, Dai LP, et al. Systematic review: exosomal microRNAs associated with pancreatic cancer for early detection and prognosis. *Eur Rev Med Pharmacol Sci* 2019;23:9351–61.
79. Kishikawa T, Otsuka M, Ohno M, Yoshikawa T, Takata A, Koike K. Circulating RNAs as new biomarkers for detecting pancreatic cancer. *World J Gastroenterol* 2015;21:8527–40.

80. Cruz-Monserrate Z, Abd-Elgalil WR, Grote T, Deng D, Ji B, Arumugam T, et al. Detection of pancreatic cancer tumours and precursor lesions by cathepsin E activity in mouse models. *Gut* 2012;61:1315–22.
81. Reader CS, Vallath S, Steele CW, Haider S, Brentnall A, Desai A, et al. The integrin alphavbeta6 drives pancreatic cancer through diverse mechanisms and represents an effective target for therapy. *J Pathol* 2019;249:332–42.
82. Steiger K, Schlitter AM, Weichert W, Esposito I, Wester HJ, Notni J. Perspective of alphavbeta6-integrin imaging for clinical management of pancreatic carcinoma and its precursor lesions. *Mol Imaging* 2017;16:1536012117709384.
83. Notni J, Reich D, Maltsev OV, Kapp TG, Steiger K, Hoffmann F, et al. In vivo PET imaging of the cancer integrin alphavbeta6 using (68)Ga-labeled cyclic RGD nonapeptides. *J Nucl Med* 2017;58:671–7.
84. Hausner SH, Bold RJ, Cheuy LY, Chew HK, Daly ME, Davis RA, et al. Preclinical development and first-in-human imaging of the integrin alphavbeta6 with [(18)F]alphavbeta6-binding peptide in metastatic carcinoma. *Clin Cancer Res* 2019;25:1206–15.
85. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer* 2013;63:11–30.
86. Fenton JJ, Weyrich MS, Durbin S, Liu Y, Bang H, Melnikow J. Prostate-specific antigen-based screening for prostate cancer: evidence report and systematic review for the US Preventive Services Task Force. *JAMA* 2018;319:1914–31.
87. Caram ME, Skolarus TA, Cooney KA. Limitations of prostate-specific antigen testing after a prostate cancer diagnosis. *Eur Urol* 2016;70:209–10.
88. Hayes JH, Barry MJ. Screening for prostate cancer with the prostate-specific antigen test: a review of current evidence. *JAMA* 2014;311:1143–9.
89. Cooperberg MR, Carroll PR. Trends in management for patients with localized prostate cancer, 1990–2013. *JAMA* 2015;314:80–2.
90. Hayes JH, Ollendorf DA, Pearson SD, et al. Active surveillance compared with initial treatment for men with low-risk prostate cancer: a decision analysis. *JAMA* 2010;304:2373–80.
91. Otero JR, Gómez BG, Ojeda JMD, Antolin AR, Cabo AV, Carlsson S, et al. Active surveillance for prostate cancer. *Int J Urol* 2016;23:211–8.
92. Lewis R, Hornberger B. Beyond the PSA test: how to better stratify a patient's risk of prostate cancer. *JAAPA* 2017;30:51–4.
93. Vickers AJ, Cronin AM, Aus G, Pihl CG, Becker C, Pettersson K, et al. Impact of recent screening on predicting the outcome of prostate cancer biopsy in men with elevated prostate-specific antigen: data from the European Randomized Study of Prostate Cancer Screening in Gothenburg, Sweden. *Cancer* 2010;116:2612–20.
94. Johnson LM, Turkbey B, Figg WD, Choyke PL. Multiparametric MRI in prostate cancer management. *Nat Rev Clin Oncol* 2014;11:346–53.
95. Rosenkrantz AB, Ginocchio LA, Cornfeld D, Froemming AT, Gupta RT, Turkbey B, et al. Interobserver reproducibility of the PI-RADS version 2 lexicon: a multicenter study of six experienced prostate radiologists. *Radiology* 2016;280:793–804.
96. Liu L, Tian Z, Zhang Z, Fei B. Computer-aided detection of prostate cancer with MRI: technology and applications. *Acad Radiol* 2016;23:1024–46.
97. Stoyanova R, Pollack A, Takhar M, Lynne C, Parra N, Lam LL, et al. Association of multiparametric MRI quantitative imaging features with prostate cancer gene expression in MRI-targeted prostate biopsies. *Oncotarget* 2016;7:53362–76.
98. Parra NA, Lu H, Choi J, Gage K, Pow-Sang J, Gillies RJ, et al. Habitats in DCE-MRI to predict clinically significant prostate cancers. *Tomography* 2019;5:68–76.
99. Tschudi Y, Pollack A, Punnen S, Ford JC, Chang YC, Soodana-Prakash N, et al. Automatic detection of prostate tumor habitats using diffusion MRI. *Sci Rep* 2018;8:16801.
100. Varghese B, Chen F, Hwang D, Palmer SL, De Castro Abreu AL, Ukimura O, et al. Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images. *Sci Rep* 2019;9:1570.
101. Smith CP, Czarniecki M, Mehralivand S, Stoyanova R, Choyke PL, Harmon S, et al. Radiomics and radiogenomics of prostate cancer. *Abdom Radiol* 2019;44:2021–9.
102. Sun Y, Reynolds HM, Parameswaran B, Wraith D, Finnegan ME, Williams S, et al. Multiparametric MRI and radiomics in prostate cancer: a review. *Australas Phys Eng Sci Med* 2019;42:3–25.
103. Garg K, Chang S, Scherzinger A. Obesity and diabetes: newer concepts in imaging. *Diabetes Technol Ther* 2013;15:351–61.
104. Li C, Li R, Zhang W. Progress in non-invasive detection of liver fibrosis. *Cancer Biol Med* 2018;15:124–36.
105. Sanyal AJ, Yoon SK, Lencioni R. The etiology of hepatocellular carcinoma and consequences for treatment. *Oncologist* 2010;15Suppl 4:14–22.
106. Knudsen ES, Gopal P, Singal AG. The changing landscape of hepatocellular carcinoma: etiology, genetics, and therapy. *Am J Pathol* 2014;184:574–83.
107. Arulanandan A, Loomba R. Non-invasive testing for NASH and NASH with advanced fibrosis: are we there yet? *Curr Hepatol Rep* 2015;14:109–18.
108. Frulio N, Trillaud H. Ultrasound elastography in liver. *Diagn Interv Imaging* 2013;94:515–34.
109. Liu D, Liu F, Xie X, Su L, Liu M, Xie X, et al. Accurate prediction of responses to transarterial chemoembolization for patients with hepatocellular carcinoma by using artificial intelligence in contrast-enhanced ultrasound. *Eur Radiol* 2020;30:2365–76.
110. Masuzaki R, Tateishi R, Yoshida H, Goto E, Sato T, Ohki T, et al. Prospective risk assessment for hepatocellular carcinoma development in patients with chronic hepatitis C by transient elastography. *Hepatology* 2009;49:1954–61.
111. Barr RG, Ferraioli G, Palmeri ML, Goodman ZD, Garcia-Tsao G, Rubin J, et al. Elastography assessment of liver fibrosis: Society of Radiologists in ultrasound consensus conference statement. *Radiology* 2015;276:845–61.
112. Venkatesh SK, Yin M, Ehman RL. Magnetic resonance elastography of liver: technique, analysis, and clinical applications. *J Magn Reson Imaging* 2013;37:544–55.
113. Tunalı I, Tan Y, Gray JE, Eschrich S, Guvenis A, Gillies R, et al. Clinical-radiomic models predict overall survival among non-small cell lung cancer patients treated with immunotherapy. *J Thorac Oncol* 2019;14:S1129.