Power System Technology
ISSN:1000-3673

Received: 09 - 08 - 2026 *Revised: 23 - 09 - 2026* *Accepted: 05 -10 - 2026*

Artificial Intelligence-Driven Medical Imaging and Predictive Analytics for Early Tumor Detection, Classification, Treatment Response Assessment, and Disease Progression Monitoring

Yeasin Arafat[1], Md Abdus Salam[2], Farhana Akter[3], Mohammad Nazmul Hossain[4], Amena Khatun Urmi[5], Sadia Afrin[6], Mantaka Rowshon[7], Md Moshiour Rahman[8], Nur Rahman[9]

[1]Doctor of Computer Science, University of the Potomac

[2]Doctor of Computer Science, University of the Potomac

[3]MS in Business Analytics, St. Francis College, Brooklyn

[4]Information Technology and Management, St. Francis College, Brooklyn

[5]Department of Physics, Eden Mohila College

[6]Information Technology and Management,.St. Francis College, Brooklyn

[7]Data Science, Master's Degree, University of the Cumberlands

[8]Business Analytics, University of New Haven

[9]Information Technology and Management, St. Francis College, Brooklyn

Corresponding Author: Yeasin Arafat

Abstract:

Cancer outcomes depend heavily on how early a tumor is found, how accurately it is characterized, and how quickly change sin its behavior are recognized during therapy. Medical imaging is the principal source of this information, yet the volume and complexity of imaging data now exceed what human readers can interpret consistently. Artificial intelligence (AI), especially deep learning and radiomics, offers a way to extract quantitative, reproducible information from images and to convert it into predictions that support clinical decisions. This article reviews the evidence for AI across four stages of the oncology pathway:earlydetection, classification, treatment response assessment, and disease progression monitoring. It then develops a compact mathematical framework for each stage, covering convolutional detection models, receiver operating characteristic (ROC) analysis, Bayesian interpretation of screening results, tumor growth and treatment response kinetics, and survival modeling with a Cox proportional hazards formulation. Worked calculations show that (a) a classifier with 94% sensitivity and 92% specificity has a positive predictive value of only 10.6% at 1% prevalence,

(b) detecting a lesion at 0.5 cm rather than 1.0 cm gives three tumor doublings of lead time,

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

(c) a 30% decrease in diameter underRECIST 1.1 corresponds to a 65.7% decrease in volume, and (d) a high-risk imaging signature with a hazard ratio of 2.5 shortens median progression-free survival from 34.7 to 13.9 months in the illustrative model. Five graphs and three tables summarize the results. The model outputs are illustrative simulations built from stated assumptions and are not patient data. The article closes with a discussion of validation, bias, explainability, and implementation barriers.

Keywords:artificial intelligence; deep learning; radiomics; medical imaging; early cancer detection; treatment response; RECIST; predictive analytics; survival modeling

Introduction

Imaging touches nearly every decision in cancer care. Screening mammography and low-dose computed tomography (CT) identify disease before symptoms appear. Magnetic resonance imaging (MRI), CT, and positron emission tomography (PET) define the extent of disease at diagnosis. Serial scans during treatment show whether a tumor is shrinking, stable, or growing, and surveillance imaging after treatment looks for recurrence. At each of these points a clinician

must interpret an image,and that interpretation has traditionally been qualitative, based on the reader's experience and on a handful of manual measurements.

This model of interpretation has three weaknesses. First, it is variable:different readers can reach different conclusions on the same scan, and the same reader can differ across days. Second, it discards most of the information in the image, because the human eye cannot quantify subtle texture, shape, or intensity patterns across thousands of voxels. Third, it does not scale. Imaging volumes continue to grow, while the number of trained radiologists and pathologists does not grow at the same pace (Hosny et al., 2018).

AI addresses these weaknesses directly. Deep neural networks learn image features from labeled examples rather than relying on hand-written rules, and they can be applied consistently to any number of scans(Litjensetal., 2017). Radiomics, thehigh-throughputextraction of quantitative features from images, treats an image as data that can be mined and linked to outcomes (Gillies et al., 2016). Predictive analytics extends this idea by combining imaging features with clinical variables to estimate risk, response, and survival. Together, these tools promise earlier detection, more precise classification, more objective response assessment, and more accurate progression forecasting (Bi et al., 2019; Topol, 2019).

The purpose of this article is twofold. The first aim is to synthesize published evidence on AI-drivenimagingacrossfourclinicaltasks:early tumor detection, tumor classification, treatment response assessment, and disease progression monitoring. The second aim is to make the underlying quantitative reasoning explicit. Many discussions of AI in oncology

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

report headline accuracy figures without showing how those figures translate in to clinical consequences. Here, each task is accompanied by a mathematical model and a worked calculation, sothatthereader can see how sensitivity, prevalence, tumor kinetics, and hazard ratios interact.

The article is organized as follows. Section 2 reviews the evidence base. Section 3 presents the mathematical framework. Section 4 gives worked calculations and illustrative model results, supported by five graphs and three tables. Section5discusses clinical implications, and Section6 addresses limitations and future directions. Section 7 concludes.

Background And Eviden Cereview

Early tumor detection

Early detection is the are a in which AI has produced some of its most striking results. Inbreast cancer screening, McKinney et al. (2020) evaluated an AI system on large datasets from the United Kingdom and the United States.

Comparedwiththeoriginalradiologistinterpretations, the system reduced false positives by 5.7% in the US dataset and 1.2% in the UK dataset, and reduced false negatives by 9.4% and 2.7%, respectively. In an independent reader study, the system outperformed all six radiologists who participated. These results matter because screening programs are judged both by the cancers they find and by the unnecessary recalls and biopsies they cause.

In lung cancer screening, Ardilaetal.(2019)trained a three-dimensionaldeeplearningmodelon low-dose chest CT from the National Lung Screening Trial. The model achieved an area under the ROC curve (AUC) of 94.4% for predicting malignancy and, when prior imaging was available, per formed at least as well as radiologists. The use of three-dimensionalconvolutionsis important, because nodules are volumetric objects whose shape and context across slices carry diagnostic information.

Pathology images show a similar pattern. Bejnordi et al. (2017) compared deep learning algorithmswith pathologistsfor detecting lymph node metastases in women with breast cancer. The best algorithm reached an AUCof0.994, where as the pathologists reviewing lidesundera simulated time constraint reached a meanAUCof0.810. Although whole-slidepathologyisnot radiology, the example shows how AI can perform well in tasks that require searching a very large image for a very small abnormality.

Tumor classification

Classification asks whether a lesion is benign or malignant, what sub type it belongs to, andwhat molecular or prognostic category it falls into. Esteva et al. (2017) trained a convolutional neural network on a dataset of roughly 129,000 clinical images and tested it against board-certified dermatologists for skin lesion classification. The network matched

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

dermatologist-level performance on the tasks tested. This study is widely cited because it showed that a general-purpose image classification architecture, adapted through transfer learning, could handle a clinically meaningful diagnostic task.

In radiology, classification increasingly involves radiomics. Gillies et al. (2016) described how features describing intensity, shape, texture, and wave let-filtered patterns can be extracted from segmented tumors and linked to phenotype, genotype, and outcome. A radiomic signature can therefore act as a non-invasive, repeatable "virtual biopsy" that captures the entire tumor, including regions that a needle biopsy might miss. Bi et al. (2019) reviewed these applications across cancer types and emphasized that the quality of segmentation, image acquisition, and feature standardization strongly influences reproducibility.

Treatment responseassessment

Treatment response in solid tumors is most often assessed with the Response Evaluation Criteria in Solid Tumours, version 1.1 (RECIST 1.1) (Eisenhauer et al., 2009). RECIST sums the longest diametersofselectedtargetlesionsandclassifieschangeintocompleteresponse, partial response (at least a 30% decrease from baseline), progressive disease (at least a 20% increase from the smallest sum on study, with a minimum absolute increase of 5 mm), or stable disease. RECIST has the advantage of being simple, standardized, and widely accepted in clinical trials.

Its limitations are equally well known. Diameter is a one-dimensional summary of a three-dimensional object, manual measurement is variable, and cytostatic or immunotherapy agents can change tumor biology without producing early shrinkage. AI can support response assessment through automated segmentation, volumetric measurement, and the extraction of early imaging biomarkers that predict later response (Hosnyetal.,2018; Bietal.,2019). Section 4 shows quantitatively how diameter-based and volume-based thresholds relate to each other.

Disease progression monitoring

During follow-up, theclinicalquestionshiftsfrom"what is this?"to"what will happen next?" Predictive analytics models combine serial imaging features, clinical variables, and some times genomic or laboratory data to estimate the probability of progression, recurrence, or death.

Survival models, suchastheCoxproportionalhazardsmodel, are the standard way to relate such features to time-to-event outcomes. AI-derived imaging features can serve as covariates in these models or can be fed to neural survival models directly (Bi et al., 2019).

Topol (2019) argued that the real test of such systems is not retrospective accuracy but prospective clinical benefit, and that the combination of human and machine judgment is likely to outperform either alone. This theme returns in the discussion.

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

Summary of reported performance

Table1summarizesthekeyresultsfromthestudiesdiscussedabove. Thefiguresarequoted from the original publications and should be interpreted within the context of each study's population, reference standard, and evaluation design.

Table1.SelectedpublishedperformanceofAIsystemsinoncologicimagingand pathology

Clinical task	Study	Modality	Reportedresult
Breast cancer screening	McKinneyet al. (2020)	Mammography	False positives reduced by 5.7% (US) and 1.2%(UK); false negatives reduced by9.4% (US) and 2.7% (UK)
Lung cancer screening	Ardilaetal. (2019)	Low-dose chest CT	AUCof94.4% for malignancy prediction
Lymph node met a stasis detection	Bejnordiet al. (2017)	Whole-slide pathology	BestalgorithmAUC0.994vs.pathologist mean AUC 0.810 (time-constrained)
Skin lesion classification	Estevaetal.(2017)	Clinical photographs, dermoscopy	Performance comparable to board-certified dermatologists

Mathematical Framework

This section defines the models used in the remainder of the article. The notation is kept as simple as possible, and each model is tied to one of the four clinical tasks.

Detection model: convolutional network with a probabilistic output

Let **x** denotean imageor image volume and y $\in\{0, 1\}$indicate the absence or presence of a tumor. A convolutional neural net work(CNN)is a composition of layers. Each convolutional layer computes feature maps by sliding a small kernel *K* over its input:

$$(K * I)(i,j) = \Sigma_{\mathrm{m}}\Sigma_{\mathrm{n}} I(i+m, j+n) \cdot K(m,n)$$

Nonlinearactivationfunctionsandpoolingoperationsfollow, andthefinallayerconvertsthe learned features into a probability using the logistic (sigmoid) function:

$$p = P(y\ = 1 \mid \boldsymbol{x}) = 1\ /(1\ + e\text{^}(-z)), where z = \ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x})\ + b$$

Here **h**(**x**) is the final feature vector, **w** is a weight vector, and *b* is a bias. The net work is trained by minimizing the binary cross-entropy loss over *N* labeled examples:

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

$L=-(1/N)\ \Sigma_i[\ y_i \ln p_i+(1-y_i)\ln(1-p_i)\ ]$

A decision threshold τturns the probability into abinary call:the case is flagged as suspicious if $p\geq\tau$. Raisingττrades sensitivity for specificity, and this trade-offisthesubjectofROCanalysis below.

For segmentation, which is needed for volumetric measurement, a common training objective is the Dice loss. If A is the predicted tumor mask and B the reference mask, the Dice similarity coefficient is

$$DSC = 2|A \cap B|/(|A| + |B|)$$

and the loss is1−DSC. ADSCof1indicates perfect over lap and0indicates none.

Diagnostic performancemetrics

Let TP, FP, TN, and FN denote true positives, false positives, true negatives, and false negatives. The standard metrics are:

Sensitivity(Se) =TP/(TP+FN)

Specificity(Sp) =TN/(TN+FP)

Accuracy=(TP+TN)/(TP+FP+TN+FN)

Positive predictive value(PPV)=TP/(TP+FP)

Negative predictive value(NPV)=TN/(TN+FN)

F1score=2TP/(2TP+FP+FN)

Positive like lihood ratio LR^+=Se/(1−Sp); negative like lihood ratio LR^-=(1−Se)/Sp

Sensitivity and specificity describe the test; PPVandNPVdescribewhataresultmeansfora patient and therefore depend on disease prevalence π. By Bayes' theorem,

$$PPV = (Se \cdot \pi)/[Se \cdot \pi + (1 - Sp)(1 - \pi)]$$

$$NPV = [Sp \cdot (1 - \pi)]\ /[Sp \cdot (1 - \pi) + (1 - Se) \cdot \pi]$$

The same result can be written in odds form: post-test odds=pre-test odds×LR^+.

ROC analysis and the binormal model

TheROCcurveplotsthetruepositiverate(TPR, equal to sensitivity)against the false positive rate (FPR, equal to 1 − specificity) as the threshold τ varies. The area under the curve (AUC) equals the probability that a randomly chosen positive case receives a higher score than a randomly chosen negative case.

TogenerateillustrativeROCcurves, we assume the classifier'sscoreisnormallydistributed with equal variance in the two classes, with means separated by d′ standard deviations (the

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

binormal, equal-variance model). Then

$$TPR = \Phi(d' - \Phi^{-1}(1 - FPR))$$

$$AUC = \Phi(d'/\sqrt{2}\,)$$

whereΦisthestandardnormalcumulativedistributionfunction. Invertingthesecondequation givesd′=√2 · Φ⁻¹(AUC). Thisallowsusto produce an ROCcurvefor any targetAUC.

Tumor growth and responsekinetics

Untreated tumor growth over short intervals is of ten approximated by exponentialgrowth:

V(t) =V₀ e^(kt), with doubling time T_d = ln 2/k

where V_0 is the initial volume and k is the growth rate constant. For a spherical tumor of diameter d, V = (π/6)d3. Because volume scales with the cube of diameter, small changes in diameter correspond to large change sin volume. This relationship is central to under standing both detection lead time and RECIST thresholds.

For tumors under treatment, atwo-compartment model is widely used. A fraction $f$of the tumor issensitiveandregressesatrate d, while the remaining fraction$(1-f)$is resistant and continues to grow at rate g:

$$V(t) = V_0\,[fe\text{^}(-dt) + (1 - f)e\text{^}(gt)\,]$$

Thisfunctiondeclinesinitially, reaches a minimum(thenadir), and then rises as the resistant population dominates. Setting dV/dt = 0 gives the nadir time:

$$t* = ln[\,d\,f\,/(g\,(1 - f))\,]\,/(d + g)$$

The model captures the clinically familiar pattern of response followed by regrowth, and it providesamathematicalbasisfordefiningprogressionrelativetothenadir, asRECIST1.1does.

Progression and survival modeling

Let T be the time to progression or death. The survival function S(t)=P(T>t) and hazard function h(t) are related by S(t) = exp(−∫₀ᵗ h(u) du). The Cox proportional hazards model expresses the hazard for a patient with covariate vector **x** as

$$h(t\,|\boldsymbol{x}) = h_0(t) \cdot exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{x})$$

so that the hazard ratio associated with a one-unit increase in covariate x_j is exp(β_j). Inour illustration, the covariate is a binary AI-derived imaging risk class (high vs. low), and the baseline hazard is taken as constant, h₀(t)=λ. ThenS(t)=exp(−λe^(βx)t)and the median survival time is ln 2 / (λ e^(βx)).

Finally, an imaging-based risks core can be formed by logistic regression on radiomic features fᵢ,

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

…, f_m:

$\text{logit}(p)=\beta_0+\Sigma_j\beta_j f_j$

with regularization (for example, an L1penalty)to limit over fitting when the number of features is large relative to the number of patients (Gillies et al., 2016).

Worked Calculations and IllustrativeResults

All numerical results in this section are produced by the models in Section3under the stated assumptions. They are simulations intended to show how the mathematics behaves, not measurements from patients.

Bayesian inter pretation of a screening result

Suppose an AI screening classifier has a sensitivity of 0.94 and a specificity of 0.92 and is applied to 10,000 people at a cancer prevalence of 1%. The expected counts follow directly. Of the10,000 people,100 have cancer and 9,900 donot. True positives are$0.94\times100=94$, which

Leaves 6 false negatives. False positives are$(1-0.92)\times9,900=792$, which leaves9,108true negatives. Table 2 turns these counts into the standard metrics.

Table2.Diagnostic metrics for the worked screening example(10,000people, prevalence1%)

Metric	Formula	Calculation	Value
Sensitivity	TP/(TP+FN)	94 /100	94.0%
Specificity	TN/(TN+FP)	9,108 /9,900	92.0%
Accuracy	(TP+TN)/N	(94 +9,108) /10,000	92.0%
Positive predictive value	TP/(TP+FP)	94 /886	10.6%
Negative predictive value	TN/(TN+FN)	9,108 /9,114	99.93%

Metric Formula	Calculation	Value
F1 score	2TP/(2TP+FP+ FN)	188 /986	0.191
Positive likelihood ratio	Se/(1 −Sp)	0.94 /0.08	11.75
Negative likelihood ratio	(1 −Se) /Sp	0.06 /0.92	0.065

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

The table exposes the central tension of screening. Accuracy is 92%, and the NPV is above 99.9%, so a negative result is highly reassuring. Yet the PPV is only 10.6%: of the 886 people flagged(94+792), only94truly have cancer, which means roughly9.4recalls for every cancer found. The odds form of Bayes' theorem gives the same answer. The pre-test odds are 0.01 /

0.99=0.0101; multiplyingbyLR^{+}=11.75gives post-testoddsof0.1187, which correspondsto

a probability of 0.1187 /1.1187 =10.6%.

PPV is therefore a property of the population as much as of the algorithm. Figure 1 recomputes PPV for the same classifier across a range of prevalence values. At 5% prevalence the PPV is 38.2%, at10% itis56.6%, andat30% it reaches83.4%. A model validated in a high-riskreferral clinic can look very different when deployed in a general screening population, which is why prevalence must be reported alongside sensitivity and specificity.

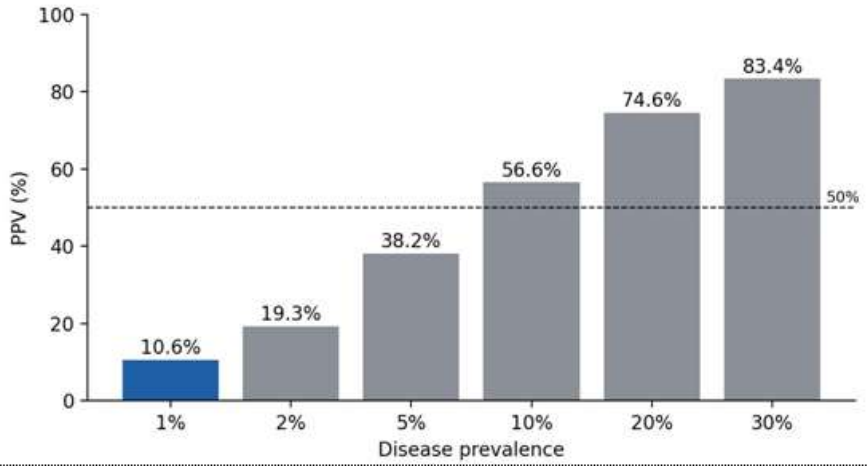


Figure1.Positive predictive value as a function of prevalence (sensitivity94%, specificity92%)

The uncertainty in the estimates also matters. With only 100 cancers, the standard error of the sensitivity is $\sqrt{(0.94 \times 0.06 / 100)} = 0.0238$, and the 95% confidence interval is $0.94 \pm 1.96 \times 0.0238$, or89.3% to98.7%. The specificity, estimated from9,900healthy people, has a standard error of $\sqrt{(0.92 \times 0.08 / 9{,}900)} = 0.0027$ and an interval of only 91.5% to 92.5%. A headline

sensitivity from a small positive cohort is therefore far less precise than the specificity,apoint that is often lost when results are summarized as a single number.

ROC analysis and the cost of a weaker classifier

Using$d' = \sqrt{2} \cdot \Phi^{-1}(\text{AUC})$, thetargetAUCvaluesof0.85,0.90, and0.95correspondtod′of 1.47, 1.81, and 2.33. Substituting into $\text{TPR} = \Phi(d' - \Phi^{-1}(1 - \text{FPR}))$ generates the curves in Figure 2. At a

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

false positive rate of 0.10, the sensitivities are 57.3%, 70.2%, and 85.2%.

The same relationship can be run in reverse to ask what a clinical target costs. To reach a sensitivity of90%, the classifier needs$\Phi^{-1}(0.90)=1.2816$, so$\Phi^{-1}(1-\mathrm{FPR})=d'-1.2816$. For AUC = 0.95 this gives 2.326 − 1.2816 = 1.0446, so 1 − FPR = 0.852 and FPR = 14.8%. For

AUC=0.85 itgives1.466−1.2816=0.1846, so 1−FPR=0.573and FPR=42.7%. In other

words,the weaker classifier produces nearly three times as many false alarms to deliver the same sensitivity. In the screening example above, that difference translates directly into recalls, biopsies, and anxiety.

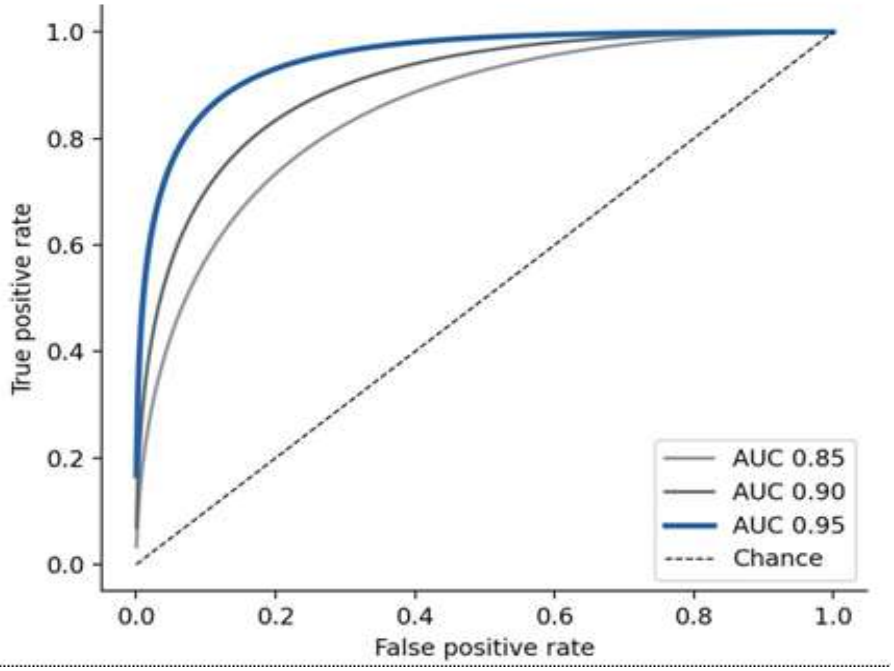


Figure2.Simulated ROC curves for three levels of discrimination (binormal model)

Detection leadtime

A more sensitive AI system is valuable partly be cause it can find smaller tumors. The size of the benefit follows from the geometry of the growth model. For a sphere, V = (π/6)d3, so a 0.5 cm lesion has a volume of (π/6)(0.125) = 0.0654 cm3 and a 1.0 cm lesion has a volume of 0.5236 cm3. The ratio is 8, which equals 23. Under exponential growth, the number of doublings

between the two sizes is $\log_2(8) = 3$, and the lead time is 3 × T_d. For a doubling time of 100 days(k=ln2/100=0.00693perday), the lead time is300days. Figure3show show this scales with the doubling time.

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

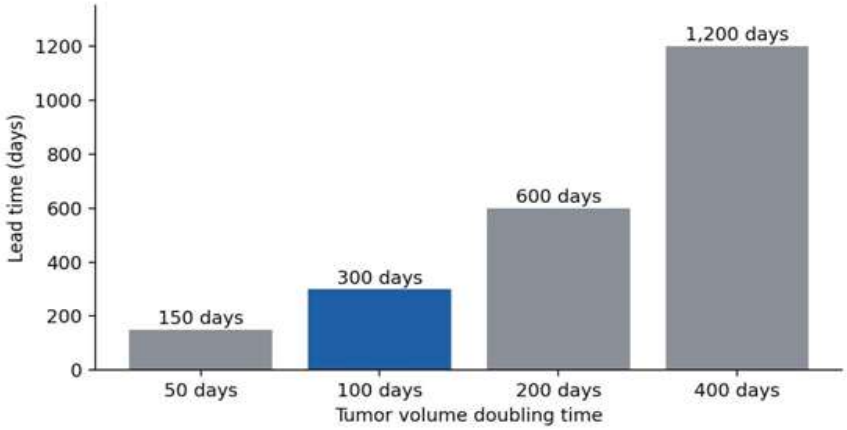


Figure3.Leadtimegainedbydetectingatumorathalfits diameter

This calculation shows why small improvements in detectable size can matter so much for slow-growing tumors, but it also invites caution. Lead time is not the same as clinical benefit. If the tumor is indolent, earlier detection may only lengthen the period during which the patient knows about it, and may contribute to overdiagnosis. Real tumors also depart from exponential growth, and growth rates vary widely between patients. The model gives an order of magnitude, not a prediction for an individual.

Treatment response: RECIST thresholds versus volume

Because volume scales with the cube of diameter, a diameter change by a factor *r*corresponds to a volume change by a factor*r*3. Table3 applies this to the RECIST1.1 thresholds (Eisenhaueret al., 2009).

Table3.Relationshipbetweendiameterchangeandvolumechangeforasphericallesion

Diameter change	Volume ratio (r^3)	Volume change	RECIST 1.1reading (single target lesion)
−30%	0.343	−65.7%	Partial response(threshold)
−20%	0.512	−48.8%	Stabledisease
−10%	0.729	−27.1%	Stabledisease
+10%	1.331	+33.1%	Stabledisease
+20%	1.728	+72.8%	Progressive disease(threshold; also require sanabsolute

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

Diameter change

Volume ratio (r^3)

Volume change

RECIST1.1reading (single target lesion)

increase of at least5mm)

+30% 2.197 +119.7% Progressive disease

A tumor must therefore lose about two-thirdsofitsvolumetoqualifyasapartialresponse, andit must gain nearly three-quarters of its volume to be called progressive. These thresholds are reasonable for manual measurement, where reader error is large, but they are coarse compared with what automated volumetry can resolve.

Thetwo-compartmentmodelillustratestheconsequences. TakeV_0=10cm3, a sensitive fraction f = 0.7 regressing at d = 0.03 per day, and a resistant fraction regrowing at g = 0.005 per day.

Thenadir timeist* = ln[(0.03 ×0.7) /(0.005 ×0.3)] /(0.03+0.005) =ln(14)/0.035 =75.4

days. At that time V = 10 × [0.7e^(−2.262) + 0.3e^(0.377)] = 5.10 cm3, a 49% reduction in volume. The corresponding diameter ratio is0.510^(1/3)=0.799, a20.1% reduction in diameter. Under RECIST 1.1 this patient is classified as having stable disease, even though the tumor volume has almost halved. An automated volumetric biomarker would register a strong early response.

Progression follows a similar logic. At the nadir the diameter is (6V/π)^(1/3) = 2.14 cm. A 20% increase would be only0.43cm, which is less than the5mmabsoluteminimum, sotheabsolute rule is the bind in gone:the diameter must reach2.64cm, which corresponds to a volume of9.59 cm3, an 88% increase over the nadir volume. Solving the model numerically, this volume is reached at about day 232. An untreated tumor growing at k = 0.007 per day would by then have reached 10 × e^(0.007 × 232) = 50.7 cm3. Figure 4 shows both trajectories. Tracking volume with AI segmentation could allow the regrowth to be recognized earlier, once the measurement error of the segmentation (for example, a Dice coefficient above 0.9 against expert contours) is shown to be smaller than the change being detected.

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

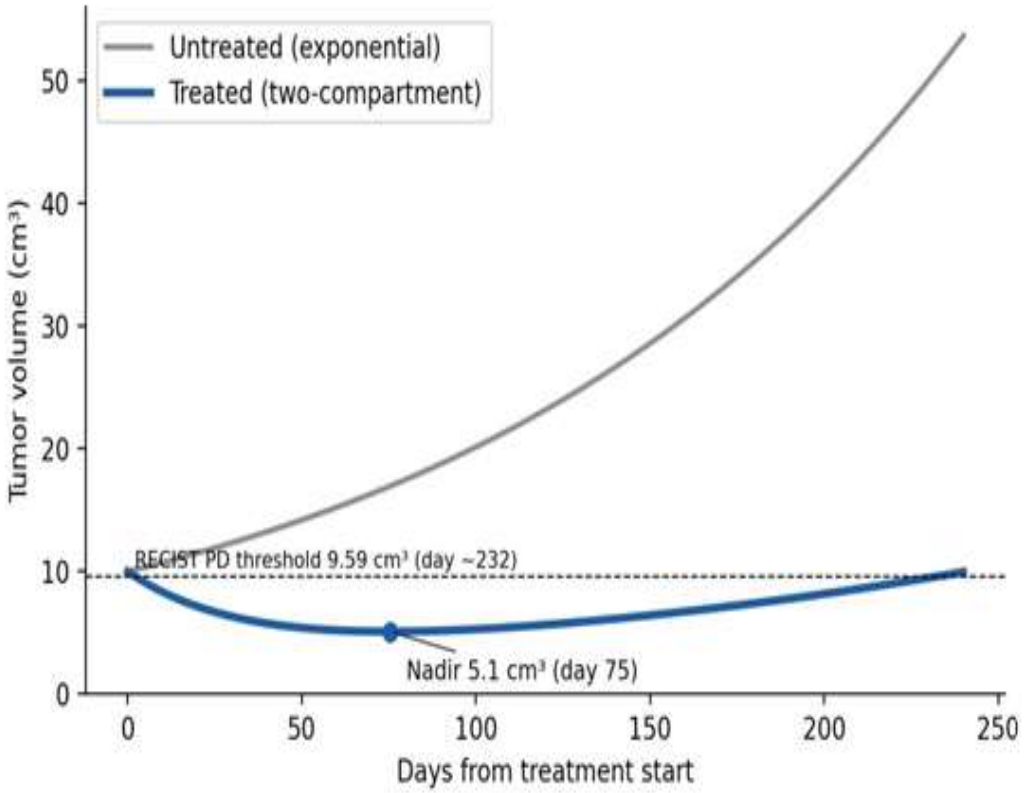


Figure 4. Simulatedtumorvolumewithandwithouttreatment

Progression monitoring and risk stratification

Suppose an AI-derived radiomics core is dichotomized in to low-risk and high-risk groups, anda Cox model estimates a coefficient $\beta = \ln(2.5) = 0.916$ for the high-risk class, which is a hazard ratio of 2.5. With a constant baseline hazard of $\lambda = 0.02$ per month, the hazard is 0.02 for the low-risk group and $0.02 \times 2.5 = 0.05$ for the high-risk group. The survival functions are $S(t) = e^{(-0.02t)}$ and $e^{(-0.05t)}$. Median progression-free survival is $\ln 2 / 0.02 = 34.7$ months for the low-risk group and $\ln 2 / 0.05 = 13.9$ months for the high-risk group. At 12 months the progression-free proportions are78.7% and54.9%, andat24months they are61.9% and30.1%, a gap of 31.8 percentage points. Figure 5 shows the full curves.

The logistic form in Section 3.5 can be used in the same way to produce a patient-level probability. If $\text{logit}(p) = -2.0 + 1.2f_1 + 0.8f_2$and a patient has standardized featurevaluesf$_1$=1.5 and f_2=0.5, then $\text{logit}(p) = -2.0 + 1.8 + 0.4 = 0.2$, and $p = 1/(1 + e^{(-0.2)}) = 55.0\%$. Apatient-level estimate of this kind is only useful if it is well calibrated, meaning that among patients assigned a probability of 55%, about 55% actually experience the event. Discrimination (AUC) and calibration are separate properties, and both need to be reported.

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

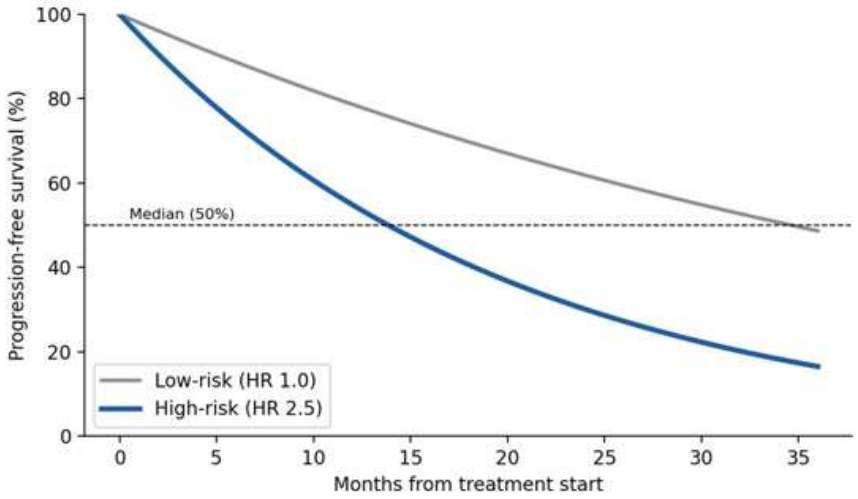


Figure5. Simulatedprogression-free survival by AI-derived riskclass

Discussion

The evidence reviewed here supports a consistent conclusion: AI can match or exceed expert performance on well-defined image interpretation tasks, particularly in screening and in the detection of small abnormalities (Ardila et al., 2019; Bejnordi et al., 2017; McKinney et al., 2020). The calculations add a second conclusion that is easy to over look. The clinical value of an AI tool depends on the setting in which it is used as much as on its accuracy. A classifier with 92% accuracy yields a PPV of only 10.6% at 1% prevalence, and the cost of an inadequate AUC is paid in false alarms at any fixed sensitivity.

For early detection, the most promising role for AI is as a second reader or triage tool that reduces both missed cancers and unnecessary recalls, which is exactly the pattern reported in breast screening (McKinney et al., 2020). The lead-time calculation shows that the upside is largest for slowly growing tumors, where as the over diagnosis risk is also high est there. Programs should therefore evaluate AI by downstream outcomes such as stage at diagnosis and interval cancers, not only by sensitivity.

For classification, radiomics and deep learning offer a route to repeatable, whole-tumor characterization (Gillies et al., 2016; Esteva et al., 2017). Their usefulness depends on standardized acquisition, robust segmentation, and external validation, since features that depend on scanner or protocol will not generalize (Bi et al., 2019). A pragmatic path is to use AI classification as decision support, presenting the probability and the image regions that drove it, rather than as an autonomous label.

For treatment response, the volume-versus-diameter analysis shows that RECIST thresholds were designed for manual measurement and are conservative when volumetry is available. A patient with a49% volume reduction can be classified asstable. AI-based volumetry and early

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

imaging biomarkers could allow earlier decisions to continue, intensify, or change therapy, but only after the new endpoints have been validated against survival or other patient-centered outcomes. Until then,

theycomplementRECISTratherthanreplaceit(Eisenhaueretal.,2009; Hosny et al., 2018).

For progression monitoring, survival models turn imaging features into time-based risk. The illustrative hazard ratio of 2.5 produces a median progression-free survival that differs by more than20months between groups, which is large enough to change surveillance intensity or trigger consideration of adjuvant therapy. Real hazard ratios from imaging signatures are typically smaller and less stable, and they must be shown to add information beyond established clinical variables.

Across all four tasks, Topol (2019)offers a useful framing: the goal is not to replace clinicians but to give them better tools, and the combination of human and machine judgment is more likely to improve outcomes than either alone.

Limitations And Futuredirections

This article has several limitations. First, the quantitative results in Section 4 come from simulations with assumed parameters. They show how the mathematics behaves, but the numbers should not be read as clinical estimates. Second, the literature summary is selective and draws on ten key sources; it is not a systematic review, and a full review would include more cancer types, modalities, and negative studies. Third, the binormal ROC model, the exponential and two-compartment growth models, and the constant-hazard Cox model are simplifications.

Real score distributions are rarely exactly normal with equal variance, tumors are heterogeneous, and hazards usually change over time.

SeveralbroaderchallengesalsoapplytoAIinoncologicimaging. Models trained on data from a few institutions may not generalize to other scanners, populations, or protocols, and they may perform unevenly across demographic groups if the training data are unrepresentative. Many networks are opaque, which complicates trust, error analysis, and regulatory review, so explainability methods and prospective performance monitoring are needed. Integration into the workflow, including reporting, audit trails, and responsibility for errors, is as important as the algorithm itself. Finally, most published evidence is retrospective; prospective randomized evaluations with patient-relevant endpoints remain scarce.

Future work should pursue four directions: prospective multi-center validation with calibration and subgroup analysis; multimodal models that combine imaging with pathology, genomics, and clinical data; longitudinal models that use sequences of scans rather than single time points; and federated learning approaches that allow model development across institutions without sharing patient data.

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 *Revised: 23 - 09 - 2026* *Accepted: 05 -10 - 2026*

Conclusion

AI-driven medical imaging and predictive analytics are changing how tumors are detected, classified, monitored during treatment, and followed over time. The published evidence shows strong performance in screening and pathology tasks, and the mathematical analysis in this article clarifies how those results translate into clinical consequences: positive predictive value depends on prevalence, a higher AUC sharply reduces false alarms at a fixed sensitivity, detecting a lesion at half its diameter can gain three tumor doublings, volumetric measurement can reveal responses that diameter criteria miss, and imaging-based risk groups can separate patientsintodistinctsurvivaltrajectories. Realizing this potential will require external validation, calibration, transparent reporting of prevalence and uncertainty, and prospective trials that measure patient outcomes. With those safeguards, AI can serve as a reliable partner to clinicians across the whole cancer care pathway.

References

[1] Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G., Naidich, D. P.,&Shetty, S.(2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nature Medicine, 25(6), 954–961. https://doi.org/10.1038/s41591-019-0447-x

[2] Bejnordi, B. E., Veta, M., vanDiest, P. J., vanGinneken, B., Karssemeijer, N., Litjens, G., van der Laak, J. A. W. M., & the CAMELYON16 Consortium. (2017). Diagnostic assessment of deeplearningalgorithmsfordetectionoflymphnodemetastasesinwomenwithbreastcancer. JAMA, 318(22), 2199–2210. https://doi.org/10.1001/jama.2017.14585

[3] Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., Allison, T.,

[4] Arnaout, O., Abbosh, C., Dunn, I. F., Mak, R. H., Tamimi, R. M., Tempany, C. M., Swanton, C.,

[5] Hoffmann, U., Schwartz, L. H., Gillies, R. J., Huang, R. Y., & Aerts, H. J. W. L. (2019). Artificial intelligence in cancer imaging:Clinical challenges and applications. CA:ACancer Journal for Clinicians, 69(2), 127–157. https://doi.org/10.3322/caac.21552

[6] Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D.,&Verweij, J.(2009). New response evaluation criteria in solid tumours:Revised RECIST guideline (version 1.1). European Journal of Cancer, 45(2), 228–247. https://doi.org/10.1016/j.ejca.2008.10.026

Power System Technology
ISSN:1000-3673
Received: 09 - 08 - 2026 Revised: 23 - 09 - 2026 Accepted: 05 -10 - 2026

[7] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature,542(7639), 115–118. https://doi.org/10.1038/nature21056

[8] Gillies, R. J., Kinahan, P. E.,&Hricak, H.(2016). Radiomics:Images are more than pictures, they are data. Radiology, 278(2), 563–577. https://doi.org/10.1148/radiol.2015151169

[9] Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H.,&Aerts, H. J. W. L.(2018). Artificial intelligence in radiology. Nature Reviews Cancer, 18(8), 500–510. https://doi.org/10.1038/s41568-018-0016-5

[10] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak,

[11] J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis.Medical Image Analysis,42,60–88.https://doi.org/10.1016/j.media.2017.07.005

[12] McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., … Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. Nature, 577(7788), 89–94. https://doi.org/10.1038/s41586-019-1799-6

[13] Topol, E. J.(2019). High-performance medicine:The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7