

RESEARCH ARTICLE

Cost-effectiveness analysis of artificial intelligence-assisted risk stratification of indeterminate pulmonary nodules

Caroline M. Godfrey¹, Ashley A. Leech², Kevin C. McGann¹, Jinyi Zhu², Hannah N. Marmor³, Sophia Pena³, Lyndsey C. Pickup⁴, Fabien Maldonado⁵, Evan C. Osmundson⁶, Stacie B. Dusetzina², Eric L. Grogan^{3,7*}, Stephen A. Deppen^{3,7}

1 Department of Surgery, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America, **2** Department of Health Policy, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America, **3** Department of Thoracic Surgery, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America, **4** Optellum Ltd, Oxford, United Kingdom, **5** Division of Allergy, Pulmonary, and Critical Care Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America, **6** Department of Radiation Oncology, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America, **7** Tennessee Valley Healthcare System, Nashville, Tennessee, United States of America

* eric.grogan@vumc.org



OPEN ACCESS

Citation: Godfrey CM, Leech AA, McGann KC, Zhu J, Marmor HN, Pena S, et al. (2026) Cost-effectiveness analysis of artificial intelligence-assisted risk stratification of indeterminate pulmonary nodules. PLoS One 21(3): e0343492. <https://doi.org/10.1371/journal.pone.0343492>

Editor: Jun Hyeok Lim, Inha University Hospital, KOREA, REPUBLIC OF

Received: July 31, 2025

Accepted: February 7, 2026

Published: March 5, 2026

Copyright: © 2026 Godfrey et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All relevant data are within the manuscript and its [Supporting Information](#) files. We have now made all of raw data available in our submission in our tables and figures, and included the full variable list file. Collectively, this provides all data required

Abstract

Background

Artificial intelligence-based radiomic approaches have been shown to accurately evaluate indeterminate pulmonary nodules. With the expansion of lung cancer screening and utilization of computed tomography imaging, indeterminate pulmonary nodules requiring diagnostic evaluation are increasingly common. Accurate non-invasive characterization may reduce time to cancer diagnosis and decrease invasive procedures for benign disease, but the cost-effectiveness of AI-based methods has not been quantified. We sought to evaluate the cost-effectiveness of AI-assisted clinician evaluation compared to clinician evaluation alone for the cancer risk stratification of patients with indeterminate pulmonary nodules.

Methods

We constructed a decision model assuming guideline-based care from a payer perspective with a lifetime horizon. The base case is a 1.1 cm incidentally discovered IPN in a 60-year-old operative candidate in a clinical population with a 65% malignancy prevalence. Cost per life-year gained (LYG) was the primary outcome. We conducted deterministic sensitivity analyses on all parameters and performed a probabilistic sensitivity analysis. Given clinical variability of malignancy prevalence, we assessed the malignancy prevalence threshold at which utilization of AI would be cost-effective.

to duplicate our work. We have additionally included our actual TreeAge software, which if future users wanted/had access to the TreeAge software could be used to run our full model immediately. This is not necessary, however, as the complete raw data and variables list is also provided -- meaning the user could run the model on any software of their choosing.

Funding: Dr. Godfrey and Dr. McGann are supported by the National Institutes of Health (T32CA106183). Dr. Leech is supported by the National Institute on Drug Abuse of the National Institutes of Health (K01DA050740). Dr. Maldonado is supported by the National Institutes of Health (R01CA253923). Dr. Dusetzina is supported by the National Institutes of Health (2P30CA068485). Dr. Grogan is supported by the National Institutes of Health (U01CA152662, R01CA252964) and was previously a recipient of the Department of VA Health Services Research and Development Service Career Development Award (10-024). Dr. Deppen is supported by the National Institutes of Health (U01CA152662). Lyndsey C. Pickup is an employee of Optellum Ltd., a commercial company which produces an artificial intelligence radiomic tool for use in indeterminate pulmonary nodules. Optellum Ltd. provided support in the form of salary for this author [LCP] but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing interests: We have read the Journal's policy, and the authors of this manuscript have the following competing interests: Lyndsey C. Pickup is an employee of Optellum Ltd., which produces an artificial intelligence radiomic tool for use in indeterminate pulmonary nodules. Optellum Ltd. provided support in the form of salary for this author [LCP] but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. This does not alter our adherence to PLOS ONE policies on sharing data and materials. The authors do not have any other relevant declarations relating to employment, consultancy, patents, products in development, or marketed products to declare.

Results

AI-supported clinician risk stratification resulted in an increase of 0.03 life years compared to clinician alone. With a 65% malignancy prevalence, AI was cost-effective with an incremental cost-effectiveness ratio (ICER) of \$4,485/LYG. When the malignancy prevalence was <5%, the ICER for AI support exceeded a standard willingness-to-pay threshold of \$100,000/LYG.

Conclusions

In clinical settings with a pre-test probability of malignancy exceeding 5%, AI-supported IPN risk stratification is cost-effective compared to clinician assessment alone.

Introduction

Nearly 15 million Americans are now eligible for annual lung cancer screening with low dose computed tomography (CT) imaging following the 2021 expansion of lung cancer screening guidelines by the United States Preventive Services Task Force [1]. In addition to screening-identified pulmonary nodules, CT is increasingly used for medical diagnostic imaging, leading to a larger number of incidentally-identified pulmonary nodules [2]. This influx of indeterminate pulmonary nodules (IPN) represents a significant healthcare burden as IPNs require diagnostic evaluation to estimate the probability of lung cancer.

Currently, an estimated 10–15% of surgical resections are performed for benign disease [3,4]. Though reducing the rate of non-therapeutic surgical resection is warranted, this must be balanced with the imperative to identify and diagnose lung cancer in its earliest, most treatable stage. Guidelines published by the American College of Chest Physicians recommend initial nodule stratification into low-probability (<5% probability of malignancy), intermediate-probability (5–65% probability of malignancy), and high-probability groups (>65% probability of malignancy) [5]. Correct classification of benign nodules into the low-probability group and malignant nodules into the high-probability group may reduce invasive diagnostic testing for benign nodules and reduce time to diagnosis for malignant nodules.

New biomarkers and clinical risk prediction models are being developed to aid in correct classification of IPNs into low- and high-probability groups. Radiomics is an emerging field which uses machine learning to model the probability of malignancy based on quantitative imaging features, usually on chest CT images.

The Lung Cancer Prediction Score (LCP Score) is part of an FDA-cleared software platform developed to aid clinicians in accurate lung nodule stratification. The underlying AI was trained and validated on a wide variety of data [6–8]. When tested in a 12-reader study, the accurate classification of lung nodules improved with LCP Scores as compared to clinicians reading alone [9].

Effective July 1st, 2022, the LCP Scores have a newly assigned common procedural terminology (CPT) code and is now eligible for reimbursement through the Centers for Medicare & Medicaid Services (CMS) at a cost range of \$600–700 [10].

However, the cost effectiveness of LCP Scores for pulmonary nodule risk stratification has not yet been studied. We sought to compare the cost-effectiveness, measured in cost per life-year, of utilization of LCP Scores as a tool to improve pulmonary nodule risk stratification to clinicians reading alone by developing a model to simulate the subsequent clinical decision-making process and resulting health outcomes.

Methods

Model overview

We constructed a decision analysis model to compare IPN stratification performed by the clinician alone to stratification performed by the clinician with AI support from the payer perspective with a lifetime horizon (Fig 1). The decision tree models subsequent care based on the risk stratification assuming guideline-based management. Outcomes included discounted life years and incremental cost-effectiveness ratios (ICERs).

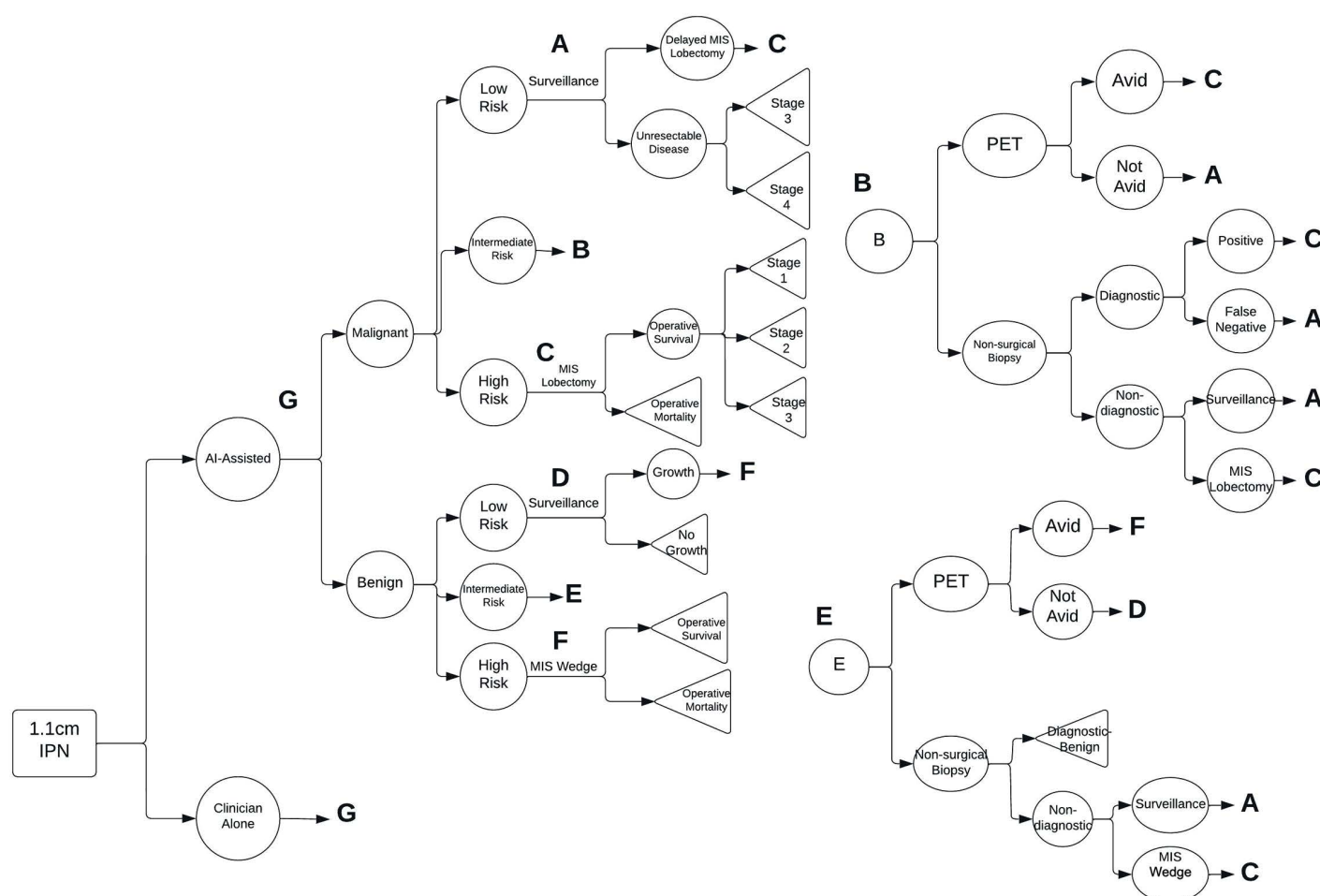


Fig 1. Decision Model Structure. The decision tree structure models the risk-stratification of an indeterminate pulmonary nodule utilizing artificial intelligence-assistance compared to the clinician alone. Repeated portions of the model have been collapsed into subtrees (A-G) for readability, each of which represents a diagnostic or management pathway that appears in various parts of the model ('A' = surveillance; 'B' = PET-CT evaluation; 'C' = minimally invasive surgical (MIS) lobectomy; 'D' = low-risk surveillance; 'E' = intermediate-risk; 'F' = MIS wedge resection; 'G' = initial risk classification).

<https://doi.org/10.1371/journal.pone.0343492.g001>

The base case is a 1.1 cm incidentally-identified IPN in a 60-year-old patient who is a medically operative candidate. The selection of the base case parameters is meant to represent a standard patient evaluated in a pulmonology or thoracic surgery clinic. The base case assumed a malignancy prevalence of 65% to represent the malignancy prevalence observed in thoracic surgery or dedicated pulmonary nodule clinics, though this parameter was varied widely as detailed in sensitivity analyses [11].

We constructed the model using TreeAge Pro Healthcare (Version 2022 R2.0) software. The model stratified the patients to low/medium/high risk, either using clinician or clinician-plus-AI reasoning, and then followed the costs and LYG related to assigning either a benign lesion or a cancer in that way. Costs and benefits (life years) were discounted at a 3% annual discount rate. We converted all costs to 2021 United States Dollars (USD) using Gross Domestic Product Price Indices [12,13].

Model structure

The model structure is a decision tree with costs and benefits assigned at each terminal node. Disease-specific costs were considered one-time costs. Unrelated medical costs accrued at each year of survival, considering the life years associated with each terminal node. We measured benefits in life years (derivation detailed in Outcomes section of Methods).

This decision tree models the initial choice of risk stratification method (AI-support versus clinician alone). The determination of the true malignancy status of the nodules (benign versus malignant) for patients is based on the malignancy prevalence within the population modeled. The decision tree subsequently incorporates probabilities of assigning patients to each cancer risk probability group based on data from the reader study that compared the clinician working alone to clinicians working with AI assistance [9]. Subsequent diagnostic steps followed clinical care consistent with the American College of Chest Physicians IPN evaluation and management guidelines [5].

Model assumptions

All high-probability nodules were assumed to be anatomically amenable to wedge resection and underwent an initial diagnostic wedge resection with frozen section analysis. Upon confirmation of a benign diagnosis, we considered patients operatively complete. Upon confirmation of a malignant diagnosis, patients then underwent a definitive lobectomy in the same operation. Though anatomic segmentectomy is now acceptable for early-stage lung cancer, given that this practice is not widespread at this time, patients with lung cancer that underwent surgical resection received a lobectomy [14].

Patients with nodules demonstrating Positron Emission Tomography (PET) avidity subsequently received a surgical biopsy, while patients with nodules that did not show PET-avidity underwent surveillance. Additionally, in the case of patients receiving a non-surgical biopsy with a non-diagnostic result, patients in the model either underwent surgical biopsy or surveillance. We varied the probability of surgical biopsy from 0 to 1 in sensitivity analyses to account for practice variation.

Model parameters

We preferentially used parameter values from meta-analyses and randomized controlled trials. If these preferred sources were not available, we used values from observational studies. When literature-derived values were not available, we substituted expert opinion or values utilized in prior cost-effectiveness analyses (Table 1). We derived risk stratification parameters from a 12-reader study previously published validating the LCP Score (Table 1) [9]. We calculated the frequency of expert reader classification into low-, intermediate-, and high-probability groups for the malignant and benign nodule groups included in the 12-reader study. The same was done for expert readers with the assistance of LCP Score. Next, we converted these values to probabilities based on the likelihood that a given malignant (or benign) nodule would be classified into each of the stated risk groups. The twelve expert readers in this study were composed of six pulmonologists

Table 1. Model Parameters.

Parameter	Base Case Parameter	Sensitivity Analysis Range	PSA distribution	References
Probability Parameters				
<i>Malignancy Prevalence</i>	0.65	.01–.8	Beta	[9]
<i>Lobectomy mortality</i>	0.022	.01–.03	Beta	[15]
<i>Wedge mortality</i>	0.0083	.005–.015	Beta	[16]
<i>Benign nodule growth</i>	0.1	.08–.12	Beta	[17]
<i>PET avid, malignant</i>	0.89	.85–.92	Beta	[18]
<i>PET avid, benign</i>	0.23	.19–.51	Beta	[18]
<i>Diagnostic biopsy, benign</i>	0.56	.45–.67	Beta	[19]
<i>Diagnostic biopsy, malignant</i>	0.92	.74–1.0	Beta	[19]
<i>False negative biopsy, malignant</i>	0.037	.03–.044	Beta	[19]
Immediate MIS resection				
<i>Stage I NSCLC</i>	0.75	.6–.82	Dirichlet	[13]
<i>Stage II NSCLC</i>	0.17	.14–.2	Dirichlet	[13]
<i>Stage III NSCLC</i>	0.08	*	Dirichlet	–
Delayed MIS resection				
<i>Stage I NSCLC, postoperative</i>	0.72	.58–.75	Dirichlet	[13]
<i>Stage II NSCLC, postoperative</i>	0.19	.15–.23	Dirichlet	[13]
<i>Stage III NSCLC, postoperative</i>	0.09	*	Dirichlet	–
<i>Progression to unresectable</i>	0.04	.02–.06	Beta	[13]
<i>Stage III NSCLC, unresectable</i>	0.5	.2–.8	Beta	[13]
<i>Stage IV NSCLC, unresectable</i>	0.5	*	Beta	–
Risk Stratification Parameters				
Clinician alone, malignant				
<i>High risk</i>	.47	.2–.58	Beta	[7]
<i>Intermediate risk</i>	.42	*		[7]
<i>Low risk</i>	.10	.01–.27	Beta	[7]
Clinician alone, benign				
<i>High risk</i>	.11	.02–.5	Beta	[7]
<i>Intermediate risk</i>	.39	*		[7]
<i>Low risk</i>	.50	.18–.71	Beta	[7]
AI-assisted, malignant				
<i>High risk</i>	.60	.45–.75	Beta	[7]
<i>Intermediate risk</i>	.35	*		[7]
<i>Low risk</i>	.05	.02–.08	Beta	[7]
AI-assisted, benign				
<i>High risk</i>	.09	.04–.15	Beta	[7]
<i>Intermediate risk</i>	.37	*		[7]
<i>Low risk</i>	.54	.36–.73	Beta	[7]
Outcome Parameters				
Life Years				
<i>Benign Disease</i>	16.44	13.15–19.73	#	[20]
<i>Stage I</i>	16.32	13.06–16.44	#	[21]
<i>Stage II</i>	8.42	6.74–10.1	#	[21]
<i>Stage III, resectable</i>	5.04	4.03–6.05	#	[21]

(Continued)

Table 1. (Continued)

Parameter	Base Case Parameter	Sensitivity Analysis Range	PSA distribution	References
Stage III, unresectable	3.84	3.07-4.61	#	[21]
Stage IV	3.34	2.67-3.84	#	[21]
Procedural Costs				
Non-Surgical Biopsy	\$2,550	\$2,040-\$3,060	Gamma	CPT 31628, 32408 DRG 200, 201
PET	\$1,480	\$1,184-\$1,776	Gamma	APC 5594
Surveillance, growth	\$163	\$130-\$196	Gamma	CPT 71250
Surveillance, stable	\$652	\$326-\$815	Gamma	CPT 71250
AI-assistance	\$650	\$100-\$1,200	Gamma	CPT 0721T
Wedge Resection	\$8,821	\$7,057-\$10,585	Gamma	DRG 166, 168
Lobectomy	\$12,944	\$10,355-\$15,533	Gamma	DRG 163, 165
Treatment Costs				
Adjuvant Chemotherapy (stage II, III) (ASP+4.3%)	\$200,446	\$160,357-\$240,535	Gamma	Table 2
Unresectable Stage III Chemotherapy (ASP+4.3%)	\$193,426	\$154,741-\$232,111	Gamma	Table 2
Stage IV Chemotherapy (ASP+4.3%)	\$421,376	\$337,101-\$505,651	Gamma	Table 2
Stage III XRT	\$166,933	\$133,546-\$200,320	Gamma	CPT 99205, 77263, 77334, 77293, 77301, 77338, 77300, 77387, 77427
Unrelated Annual Medical Costs				
Age < 65	\$9,056	\$7,245-\$10,867	Gamma	[22]
Age > 65	\$12,961	\$10,369-\$15,553	Gamma	[22]

*constrained by other model parameters; #not varied in probabilistic sensitivity analysis due to high confidence in value and stability of model with one-way sensitivity analyses; PSA: probabilistic sensitivity analysis; PET: Positron Emission Tomography; ASP: average sales price.

<https://doi.org/10.1371/journal.pone.0343492.t001>

(two with expertise in thoracic oncology) and 6 radiologists (two with expertise in thoracic radiology) described in further detail in the original reader study [9].

All malignant nodules that underwent surveillance were assumed to eventually grow on surveillance imaging and progress to lobectomy. Malignant nodules undergoing surveillance could experience a potential stage shift due to the delay in definitive management. Previously published models that investigated outcomes from delayed resection formed the basis for the stage shifting parameters [23].

For intermediate probability nodules, guidelines suggest further diagnostic evaluation via PET scan or non-surgical biopsy based on clinician and patient preference [5]. Following a non-diagnostic biopsy, patients in the model proceeded to either a diagnostic wedge resection or surveillance.

Outcomes

The primary outcomes included discounted life years and the incremental cost-effectiveness ratio (ICER) of AI-supported nodule risk stratification compared to stratification performed by the clinician alone, calculated as cost per life year gained (LYG). We used a willingness to pay (WTP) threshold of \$100,000/LYG and calculated life years associated with each outcome using the Declining Exponential Approximation of Life Expectancy (Table 1) [20, 21, 24]. For malignant diagnoses, we used life tables from the International Association for the Study of Lung Cancer Staging Project to calculate 5-year mortality rates associated with pathologic stage I, II, and III as well as clinical stage III and IV (for unresectable disease) [25]. Expected life years for the base case of patients with benign disease are from the life tables published in the National Vital Statistics Report [26]. For patients who suffered operative mortality, we assigned a life year value of zero.

Costs reflected a payer perspective using Centers for Medicare & Medicaid Services (CMS) reimbursement rates (Table 1) [22]. For procedures associated with an inpatient stay (wedge resection, lobectomy), we used diagnosis related groups (DRG) to capture all costs associated with the procedure [27]. For outpatient procedures, we estimated costs using CPT or ambulatory payment classification (APC) codes [22]. We obtained unrelated healthcare costs for each year of survival from the Medical Expenditure Panel Survey from the Agency for Healthcare Research and Quality [28].

As the impact of procedural morbidity on overall quality-adjusted life years would be minimal and temporary compared to survival differences for lung cancer, we incorporated procedural morbidity into the model via a weighted cost adjustment for the associated procedures (for details on cost weighting, see supplementary methods). We assumed the cost of procedural mortality was equivalent to the cost of the procedure with a major complication.

Anticancer medication costs are based on standard treatment regimens by stage outlined in the National Comprehensive Cancer Network (NCCN) guidelines for adenocarcinoma [29]. As lung cancer therapy is complex and becoming more highly individualized, we made assumptions in order to simply and accurately model costs of medical and radiation oncologic therapy by stage. In determining therapy costs, we considered malignant nodules as adenocarcinoma and eligible for anticancer therapy (chemotherapy, targeted therapy and/or immunotherapy) without targetable mutations. In calculating medication dosing, the base case represented the size of an average American male (200 pounds, 70 inches tall, with a body surface area of 2 m²). Additionally, while we understand that neoadjuvant therapy is now approved by the Food and Drug Administration and beginning to be incorporated in treatment for Stage IB to IIIA cancers at large tertiary centers, because its use is not yet widespread, we only included adjuvant anticancer treatment regimens [18]. Cost values for anticancer therapy are based on average sales price data plus a 4.3% markup that is standard for Medicare drug reimbursement [30]. Anticancer treatment regimens used for cost calculations by stage are outlined in Table 2 and were developed with thoracic medical oncologists in accordance with the NCCN guidelines to represent a standard regimen.

In our model, clinicians administered radiation therapy exclusively for unresectable stage III tumors in combination with anticancer therapy, followed by consolidative durvalumab treatment. We calculated this cost using the sum of CPT codes necessary to perform intensity-modulated radiotherapy (IMRT). The base case was a medically operative candidate with a lesion anatomically amenable to resection. Therefore, if the malignancy qualified for resection from an oncologic

Table 2. Chemotherapeutic Regimen by Stage.

	Agent	Dose	Number of Infusions	Total Cost
Stage I	No adjuvant therapy	—	—	\$0
Stage II-IIIa	Cisplatin	75 mg/m ²	4	\$101
	Pemetrexed	500 mg/m ²	4	\$8,617
	Pembrolizumab	200 mg	17	\$188,418
	Total			\$192,182
Unresectable Stage III	Carboplatin	AUC 2	6	\$71
	Paclitaxel	45 mg/m ²	6	\$92
	Durvalumab	10 mg/kg	26	\$185,289
	Total			\$185,452
Stage IV				
PD-L1 ≥ 50%	Pembrolizumab	200 mg	35	\$387,919
	Total			\$387,919
PD-L1 < 50%	Carboplatin	AUC 5	4	\$119
	Pemetrexed	500 mg/m ²	35	\$32,050
	Pembrolizumab	200 mg	35	\$387,919
	Total			\$420,088

<https://doi.org/10.1371/journal.pone.0343492.t002>

standpoint, the patient would proceed with resection. Consequently, stereotactic body radiation therapy (SBRT) was not included in the model.

The model did not incorporate rates of local or distant disease recurrence after initial therapeutic failure (progression to stage IV disease after concurrent chemoradiotherapy for stage III disease), and hence any costs associated with treatment for recurrence after primary therapy were not accounted for.

Sensitivity analysis

We performed a one-way sensitivity analysis on each parameter in the model, varying parameters by a minimum of $\pm 20\%$ from the base value. If the associated literature review demonstrated values with greater than $\pm 20\%$ variation, we incorporated the lower and upper bounds reported in the literature (Table 1). We varied all clinical decision probability events across a range of 0–1 to assess the model's robustness under a wide range of clinical practice conditions. As FDG-PET has demonstrated decreased specificity in regions of endemic fungal lung disease, we conducted a two-way sensitivity analysis on the frequency of obtaining FDG-PET for intermediate-probability nodules and PET specificity [31]. We varied the frequency of PET use for intermediate-probability nodules widely from 0% to 100% to represent potential practice pattern variations and varied the specificity of PET from 49% to 81% to simulate both regions that are endemic and non-endemic, respectively, for fungal lung disease [31].

We conducted a probabilistic sensitivity analysis using second-order Monte-Carlo simulations by drawing 400 random parameter sets from prespecified distributions. We defined beta distributions for all probabilities and gamma distributions for all costs. We parameterized the mean of each distribution using the base case value and setting the standard deviation as half the range of the tested parameter in deterministic sensitivity analyses. This ensured that the entire parameter range tested in deterministic sensitivity analyses fell within two standard deviations. We evaluated the probability of AI-assistance being cost-effective using cost-effectiveness acceptability curves, varying the WTP between \$0 and \$200,000/LYG.

Results

For the base case scenario, AI-supported pulmonary nodule risk stratification resulted in an increase of 0.03 life years as compared to clinician alone and an incremental cost of \$114. With a 65% pre-test probability of malignancy, AI was cost-effective at a willingness to pay threshold (WTP) of \$100,000/LYG with an incremental cost-effectiveness ratio (ICER) of \$4,485/LYG (Table 3).

Sensitivity analyses

A tornado diagram of the four most influential variables (malignancy prevalence, cost of AI, AI-assisted accuracy, and clinician accuracy) from the one-way sensitivity analyses conducted is shown in Fig 2. Notably, as the pre-test probability of malignancy was decreased, the ICER for AI-support increased and eventually exceeded the WTP threshold of \$100,000/LYG when malignancy prevalence decreased below 5% (Fig 3). When a more conservative WTP threshold of \$50,000/LYG was applied, AI-support remained cost-effective when malignancy probability was 15% or higher.

Table 3. Incremental Cost and Effectiveness of Clinician Compared to AI-Assisted Pulmonary Nodule Risk Stratification.

Strategy	Cost (2021 USD)	Incremental Cost	Effectiveness (Life Years)	Incremental Effectiveness	ICER (\$/life year)
Clinician	\$216,093	–	14.63	–	
AI-Supported	\$216,207	\$114	14.65	.03*	4,485

* Please note the incremental effectiveness represents unrounded model outputs and is .03, while the reported effectiveness (14.63 vs. 14.65 Life Years) is rounded to two decimals by convention.

<https://doi.org/10.1371/journal.pone.0343492.t003>

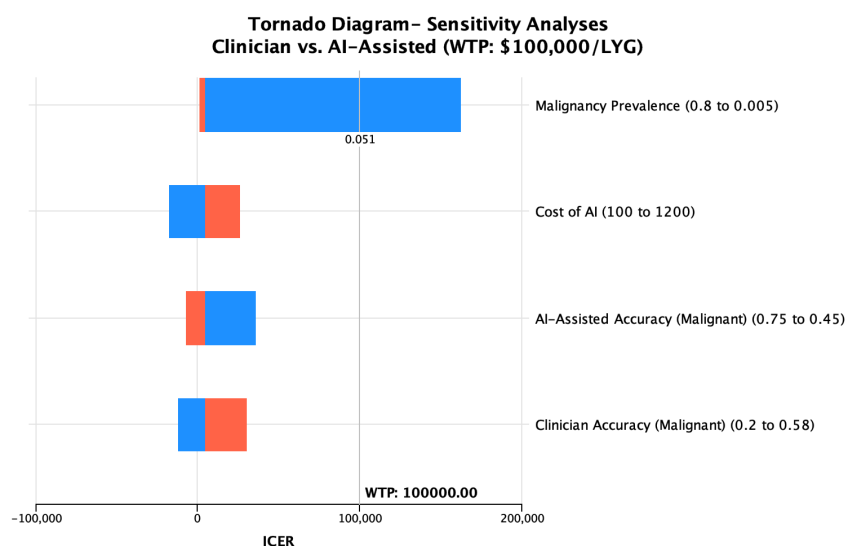


Fig 2. Tornado Diagram of Incremental Cost-Effectiveness Ratio Range of Four Most Influential Model Parameters. The tornado diagram demonstrates the incremental cost-effectiveness ratio (ICER) range for the 4 most influential parameters in one-way sensitivity analyses. Blue indicates the lower end of the varied parameter range and orange indicates the higher end of the range. The ICER exceeds the willingness-to-pay threshold only for malignancy prevalence when the prevalence is below 0.051. All other parameters remain cost-effective over the full tested range.

<https://doi.org/10.1371/journal.pone.0343492.g002>

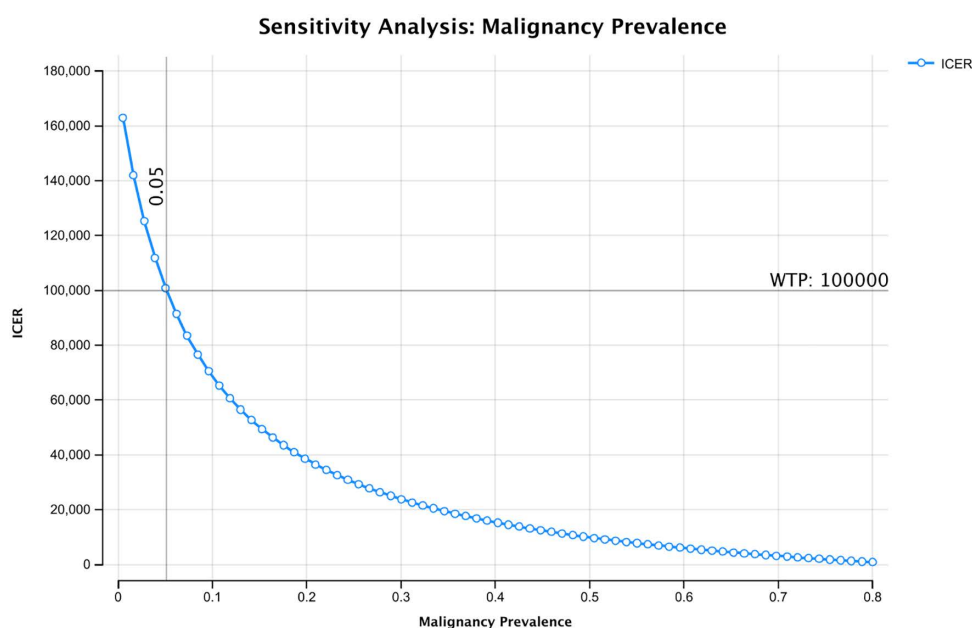


Fig 3. One-Way Sensitivity Analysis of Malignancy Prevalence. The x-axis shows the population malignancy prevalence, ranging from 0% on the left to 80% (0.8) on the right. The y-axis shows the incremental cost-effectiveness ratio (ICER) at the corresponding malignancy prevalence. The ICER exceeds the willingness-to-pay threshold when malignancy prevalence of the population is less than 5%.

<https://doi.org/10.1371/journal.pone.0343492.g003>

When varying clinician accuracy and holding AI-accuracy constant, AI-support became less cost-effective as clinician accuracy increased and eventually exceeded the WTP threshold of \$100,000/LYG when the clinician correctly identified malignant nodules as high risk greater than 65% of the time (Fig 4). The reader study data used for performance parameters shows clinician accuracy of 47% in classifying malignant nodules as high risk [9].

AI-support was increasingly cost-effective as PET specificity decreased, though remained cost-effective at the WTP threshold of \$100,000/LYG at all value combinations analyzed in the two-way sensitivity analysis.

In the probabilistic sensitivity analysis, the cost-effectiveness acceptability curves demonstrated that AI-assistance was cost-effective in greater than 50% of the 400 iterations performed when the WTP threshold was greater than \$4,058 (Fig 5). AI-assistance had a 71% probability of being cost-effective at the base-case WTP threshold (\$100,000/LYG).

Discussion

As AI is increasingly woven into medical decision-making and these platforms come to market, it is important to evaluate their cost-effectiveness. This study suggests that AI-supported IPN risk stratification, at the observed accuracies from the previously published 12-reader LCP Score study, may be cost-effective in populations with an underlying malignancy rate of 5% or greater [9]. Reported malignancy rates amongst nodules detected in a screening population are estimated around 5% [32]. Among all incidentally-discovered pulmonary nodules, malignancy prevalence is reported to be only 3%, however incidentally discovered nodules in patients with a history of tobacco use have a higher prevalence, reported at 5.5%, and nodules measuring larger than 1 cm have a 12–15% malignancy risk [33]. Malignancy prevalence in dedicated pulmonary nodule and thoracic surgery clinics is reported to be even higher, approaching 60–80% [11]. In this context, AI support would be an appropriate and cost-effective method for nodule risk stratification, though further investigation of the potential cost-effectiveness in a general screening population with malignancy prevalence near the threshold of 5% is warranted due to the greater risk of false positive findings. Additionally, real-world guideline-driven practices and radiologist expertise for thoracic screening studies vary worldwide which may also limit the potential benefit of these tools. Moreover,

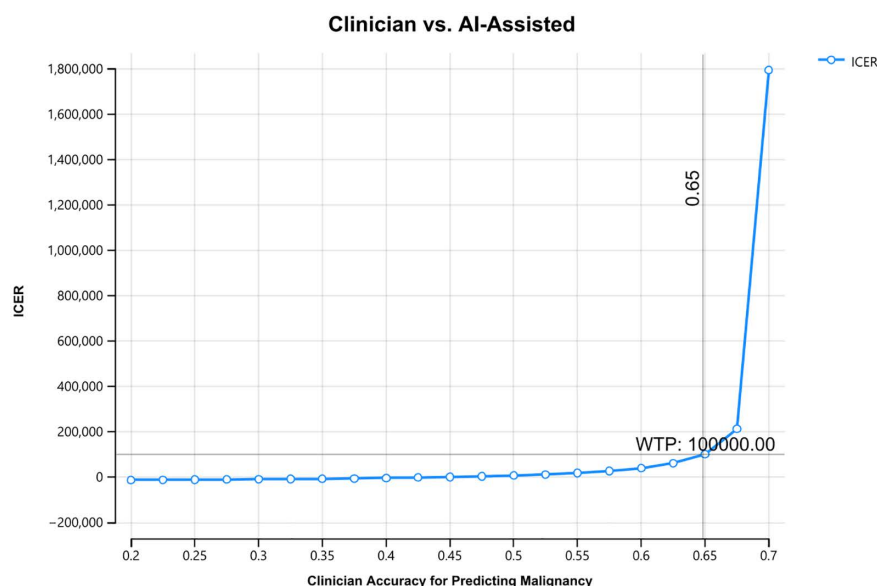


Fig 4. Sensitivity Analysis of Clinician Accuracy for Predicting Malignancy. The x-axis shows the probability a clinician correctly classifies a malignant nodule as high-risk. The y-axis shows the incremental cost-effectiveness ratio (ICER) corresponding to that probability. The ICER exceeds the willingness-to-pay threshold when the clinician accuracy for correctly identifying malignant nodules as high-risk exceeds 65%.

<https://doi.org/10.1371/journal.pone.0343492.g004>

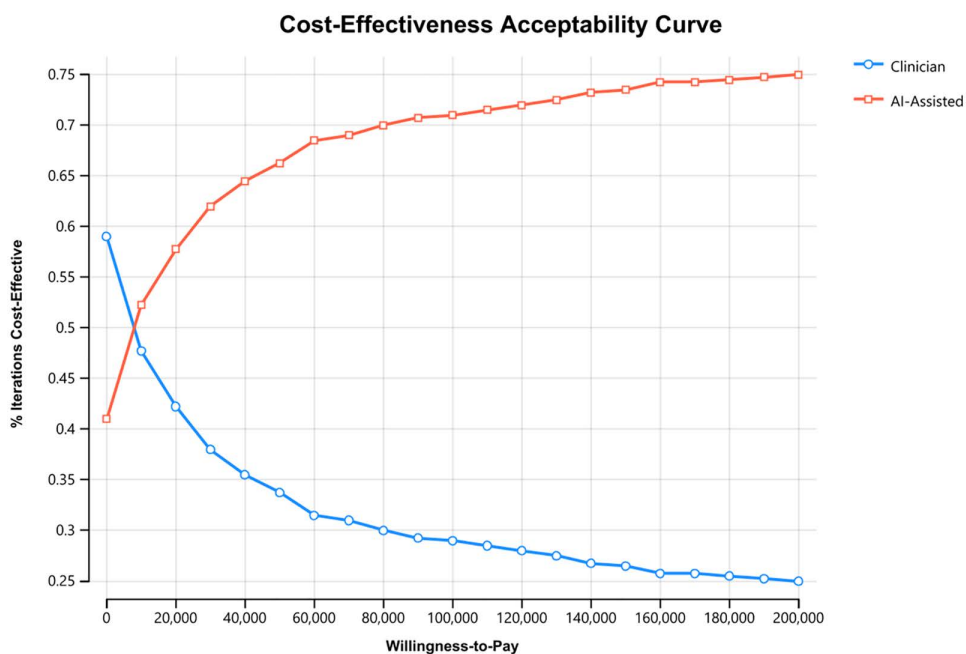


Fig 5. Probabilistic Sensitivity Analysis: Cost-Effectiveness Acceptability Curve. The x-axis represents varying willingness-to-pay thresholds. The y-axis represents the percentage of iterations of the model that resulted in either strategy being cost-effective at the corresponding willingness-to-pay. The curves cross at \$4,058, the willingness-to-pay threshold where 50% of the iterations were cost-effective for each strategy. At the base case willingness-to-pay of \$100,000/life year gained, AI-assistance is cost-effective in 71% of iterations.

<https://doi.org/10.1371/journal.pone.0343492.g005>

the racial and ethnic makeup of global populations may not be fully represented in the training cohorts used to build this, and other, AI tools.

Though this model was robustly tested in both deterministic and probabilistic sensitivity analyses which all returned similar results, supporting the conclusion of the cost-effectiveness of AI support in pulmonary nodule risk estimation, this study has limitations. One limitation of the study is that we assume guideline-based diagnostic evaluation and management and therefore the model is not reflective of true clinical practice at this time. Significant variation in management of indeterminate pulmonary nodules exists and clinical practice has been demonstrated to deviate significantly from guideline-based care [34]. This model was built to resemble guideline-based care as closely as possible. We acknowledge that real clinical practice patterns vary, particularly as guidelines and evidence in lung cancer treatment and diagnosis continue to evolve [35]. For example, emerging data demonstrating the oncologic equivalence of sublobar resections for early-stage lung cancers is likely to shift practice away from routine lobectomy for lung cancer [36]. Moreover, the surveillance branch in our model is limited in that it did not include potential nonoperative management, competing mortality, or loss to follow-up. Our base case patient was a good operative candidate and the model was intentionally structured to focus on differences in diagnostic workup as opposed to variable long-term treatment adherence. For this reason, malignant nodules detected during surveillance were assumed to eventually undergo definitive surgical management. Ultimately, it is unknown at this time how clinicians would interact with AI-based support tools and how this might affect their clinical decision-making. The cost-effectiveness, or effectiveness at all, of AI-supported risk stratification hinges on a resulting change in clinical management. If no changes occur in clinical management based on improvements in risk stratification, then the accuracy of the risk estimation is inconsequential to patient outcomes. This practice pattern variation would be best evaluated in a future pragmatic trial.

An additional limitation is that costs associated with treatment for recurrence after primary therapy were not incorporated into the model due to the high variation of these costs. As such, the model likely underestimates the cost of delayed diagnosis of early-stage malignancy due to inaccurate risk estimation. Incorporating these downstream costs is likely to make the use of AI support increasingly cost-effective due to increased cost penalization for late-stage disease treatment. Additionally, it should be noted that this model was created from the payer perspective, not the institutional perspective, and as such, does not include costs related to an institutional contract for use of an AI platform.

The cost-effectiveness of AI-supported risk stratification is heavily dependent on both clinician and AI accuracy. Another limitation of this study is that the accuracy of both the clinician and the AI-supported stratification are based on a single study based on 12 readers analyzing 300 nodules [9]. This cohort was not a consecutive clinical cohort and was cancer-enriched, albeit not compared to the 65% prevalence assumed here. A direct comparison of clinician alone and AI-supported nodule stratification in a consecutive clinical cohort is not currently available. It is possible that both clinician and AI performance would differ in varying clinical settings, but at this time, it is challenging to know whether their accuracies would change relative to each other.

In settings with particularly high clinician accuracy in lung nodule risk stratification, the cost-effectiveness of AI-support decreases as the accuracy of clinician alone and AI-supported stratification converge. Conversely, in clinical settings with lower clinician accuracy, the cost-effectiveness of AI-supported risk stratification would increase as the accuracies of the two methods diverge. Therefore, even when applied to the same patient population, the cost-effectiveness of AI-support will vary by institution according to clinician accuracy. Reports of clinician accuracy in identifying cancerous nodules are limited, making it challenging to distinguish which clinical settings are appropriate for the cost-effective utilization of AI. In the reader study used for parameterization of risk stratification in this model, the clinicians accurately identified malignant nodules as high cancer probability 47% of the time, however, another study conducted by Tanner and colleagues demonstrates a higher rate of malignant nodule identification among thoracic surgeons and pulmonologists [9,37]. Though there are limited data on this topic, in settings with limited access to sub-specialty trained pulmonologists and thoracic surgeons, AI is likely to accelerate the process of referral of patients with likely malignant pathology to a more specialized center.

Conclusion

AI-assisted pulmonary nodule risk estimation is cost-effective in most common clinical settings where the malignancy prevalence is at least 5%. With high rates of lung resections performed on benign disease and the necessity of rapid identification of lung cancer to initiate treatment in its earliest, most treatable phase, the need for accurate lung cancer risk estimation has never been greater. AI offers promise of improved risk estimation of lung nodules at reasonable cost and if used in a more automated process, could reduce some of the clinical burden on physicians to manually perform this risk stratification as IPNs are increasingly identified. A clinical trial should be considered to test the risk estimation improvement of AI over the clinician more robustly in clinical settings with varying nodule size, nodule shape, nodule location, cancer prevalence, and rates of fungal lung disease. As seen in other areas of medicine, AI cannot replace the clinician, but AI-assistance could improve their ability to evaluate and treat their patients in a cost-effective manner.

Supporting information

S1 File. Full variable definitions – Excel file. Excel file containing the complete list of variable definitions included in the model.
(XLSX)

Acknowledgments

The authors would like to thank thoracic medical oncologists, Christine Lovly, MD, PhD and Wade Iams, MD, MSCI, for their contributions to ensuring accuracy of aspects of the manuscript relating to medical oncology.

Author contributions

Conceptualization: Caroline M. Godfrey, Kevin C. McGann, Hannah N. Marmor, Lyndsey C. Pickup, Fabien Maldonado, Stacie B. Dusetzina, Eric L. Grogan, Stephen A. Deppen.

Data curation: Caroline M. Godfrey, Sophia Pena, Lyndsey C. Pickup, Evan C. Osmundson, Stephen A. Deppen.

Formal analysis: Caroline M. Godfrey, Stacie B. Dusetzina, Stephen A. Deppen.

Funding acquisition: Eric L. Grogan, Stephen A. Deppen.

Investigation: Caroline M. Godfrey, Ashley A. Leech, Stephen A. Deppen.

Methodology: Caroline M. Godfrey, Ashley A. Leech, Kevin C. McGann, Jinyi Zhu, Hannah N. Marmor, Lyndsey C. Pickup, Fabien Maldonado, Stacie B. Dusetzina, Eric L. Grogan.

Project administration: Kevin C. McGann.

Resources: Kevin C. McGann.

Supervision: Fabien Maldonado, Evan C. Osmundson, Eric L. Grogan, Stephen A. Deppen.

Validation: Ashley A. Leech, Jinyi Zhu, Stacie B. Dusetzina.

Visualization: Caroline M. Godfrey, Ashley A. Leech, Kevin C. McGann, Jinyi Zhu, Lyndsey C. Pickup.

Writing – original draft: Caroline M. Godfrey, Ashley A. Leech, Jinyi Zhu, Hannah N. Marmor, Sophia Pena, Lyndsey C. Pickup, Fabien Maldonado, Stacie B. Dusetzina, Eric L. Grogan, Stephen A. Deppen.

Writing – review & editing: Caroline M. Godfrey, Ashley A. Leech, Kevin C. McGann, Evan C. Osmundson.

References

1. Pinsky PF, Lau YK, Doubeni CA. Potential disparities by sex and race or ethnicity in lung cancer screening eligibility rates. *Chest*. 2021;160(1):341–50. <https://doi.org/10.1016/j.chest.2021.01.070> PMID: [33545164](#)
2. Gould MK, Tang T, Liu I-LA, Lee J, Zheng C, Danforth KN, et al. Recent Trends in the Identification of Incidental Pulmonary Nodules. *Am J Respir Crit Care Med*. 2015;192(10):1208–14. <https://doi.org/10.1164/rccm.201505-0990OC> PMID: [26214244](#)
3. Kuo E, Bharat A, Bontumasi N, Sanchez C, Zoole JB, Patterson GA, et al. Impact of video-assisted thoracoscopic surgery on benign resections for solitary pulmonary nodules. *Ann Thorac Surg*. 2012;93(1):266–72; <https://doi.org/10.1016/j.athoracsur.2011.08.035> PMID: [22075217](#)
4. Veronesi G, Bellomi M, Scanagatta P, Preda L, Rampinelli C, Guarize J, et al. Difficulties encountered managing nodules detected during a computed tomography lung cancer screening program. *J Thorac Cardiovasc Surg*. 2008;136(3):611–7. <https://doi.org/10.1016/j.jtcvs.2008.02.082> PMID: [18805261](#)
5. Gould MK, Donington J, Lynch WR, Mazzone PJ, Midthun DE, Naidich DP, et al. Evaluation of individuals with pulmonary nodules: When is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2013;143(5 Suppl):e93S–e120S. <https://doi.org/10.1378/chest.12-2351> PMID: [23649456](#)
6. Massion PP, Antic S, Ather S, Arteta C, Brabec J, Chen H, et al. Assessing the accuracy of a deep learning method to risk stratify indeterminate pulmonary nodules. *Am J Respir Crit Care Med*. 2020;202(2):241–9. <https://doi.org/10.1164/rccm.201903-0505OC> PMID: [32326730](#)
7. Baldwin DR, Gustafson J, Pickup L, Arteta C, Novotny P, Declerck J, et al. External validation of a convolutional neural network artificial intelligence tool to predict malignancy in pulmonary nodules. *Thorax*. 2020;75(4):306–12. <https://doi.org/10.1136/thoraxjnl-2019-214104> PMID: [32139611](#)
8. Heuvelmans MA, van Ooijen PMA, Ather S, Silva CF, Han D, Heussel CP, et al. Lung cancer prediction by Deep Learning to identify benign lung nodules. *Lung Cancer*. 2021;154:1–4. <https://doi.org/10.1016/j.lungcan.2021.01.027> PMID: [33556604](#)
9. Kim RY, Oke JL, Pickup LC, Munden RF, Dotson TL, Bellinger CR, et al. Artificial intelligence tool for assessment of indeterminate pulmonary nodules detected with CT. *Radiology*. 2022;304(3):683–91. <https://doi.org/10.1148/radiol.212182> PMID: [35608444](#)
10. Medicare Learning Network and Center for Medicare & Medicaid Services. July 2022 Update of the Hospital Outpatient Prospective Payment System (OPPS). 2022. <https://www.cms.gov/files/document/mm12761-july-2022-update-hospital-outpatient-prospective-payment-system-opps.pdf>
11. Godfrey CM, Shipe ME, Welty VF, Maiga AW, Aldrich MC, Montgomery C, et al. The Thoracic research evaluation and treatment 2.0 model: A lung cancer prediction model for indeterminate nodules referred for specialist evaluation. *Chest*. 2023;164(5):1305–14. <https://doi.org/10.1016/j.chest.2023.06.009> PMID: [37421973](#)
12. Dunn A, Grosse SD, Zuvekas SH. Adjusting health expenditures for inflation: A review of measures for health services research in the United States. *Health Serv Res*. 2018;53(1):175–96. <https://doi.org/10.1111/1475-6773.12612> PMID: [27873305](#)

13. Agency for Healthcare Research and Quality. Using Appropriate Price Indices for Analyses of Health Care Expenditures or Income Across Multiple Years. 2023.
14. Saji H, Okada M, Tsuboi M, Nakajima R, Suzuki K, Aokage K, et al. Segmentectomy versus lobectomy in small-sized peripheral non-small-cell lung cancer (JCOG0802/WJOG4607L): A multicentre, open-label, phase 3, randomised, controlled, non-inferiority trial. *Lancet*. 2022;399(10335):1607–17. [https://doi.org/10.1016/S0140-6736\(21\)02333-3](https://doi.org/10.1016/S0140-6736(21)02333-3) PMID: [35461558](#)
15. Kozower BD, Sheng S, O'Brien SM, Liptay MJ, Lau CL, Jones DR, et al. STS database risk models: Predictors of mortality and major morbidity for lung cancer resection. *Ann Thorac Surg*. 2010;90(3):875–81; discussion 881–3. <https://doi.org/10.1016/j.athoracsur.2010.03.115> PMID: [20732512](#)
16. Linden PA, D'Amico TA, Perry Y, et al. Quantifying the safety benefits of wedge resection: A society of thoracic surgery database propensity-matched analysis. *Ann Thorac Surg*. Nov 2014;98(5):1705–11. <https://doi.org/10.1016/j.athoracsur.2014.06.017>
17. Deppen SA, Davis WT, Green EA, Rickman O, Aldrich MC, Fletcher S, et al. Cost-effectiveness of initial diagnostic strategies for pulmonary nodules presenting to thoracic surgeons. *Ann Thorac Surg*. 2014;98(4):1214–22. <https://doi.org/10.1016/j.athoracsur.2014.05.025> PMID: [25087933](#)
18. Forde PM, Spicer J, Lu S, Provencio M, Mitsudomi T, Awad MM, et al. Neoadjuvant nivolumab plus chemotherapy in resectable lung cancer. *N Engl J Med*. 2022;386(21):1973–85. <https://doi.org/10.1056/NEJMoa2202170> PMID: [35403841](#)
19. Gould MK, Sanders GD, Barnett PG, Rydzak CE, Maclean CC, McClellan MB, et al. Cost-effectiveness of alternative management strategies for patients with solitary pulmonary nodules. *Ann Intern Med*. 2003;138(9):724–35. <https://doi.org/10.7326/0003-4819-138-9-200305060-00009> PMID: [12729427](#)
20. Beck JR, Pauker SG, Gottlieb JE, Klein K, Kassirer JP. A convenient approximation of life expectancy (the "DEALE"). II. Use in medical decision-making. *Am J Med*. 1982;73(6):889–97. [https://doi.org/10.1016/0002-9343\(82\)90787-2](https://doi.org/10.1016/0002-9343(82)90787-2) PMID: [6983298](#)
21. Beck JR, Kassirer JP, Pauker SG. A convenient approximation of life expectancy (the "DEALE"). I. Validation of the method. *Am J Med*. 1982;73(6):883–8. [https://doi.org/10.1016/0002-9343\(82\)90786-0](https://doi.org/10.1016/0002-9343(82)90786-0) PMID: [7148881](#)
22. Centers for Medicare & Medicaid Services. Physician Fee Schedule. 2023.
23. Shipe ME, Haddad DN, Deppen SA, Kozower BD, Grogan EL. Modeling the impact of delaying the diagnosis of non-small cell lung cancer during COVID-19. *Ann Thorac Surg*. 2021;112(1):248–54. <https://doi.org/10.1016/j.athoracsur.2020.08.025> PMID: [33091367](#)
24. Vanness DJ, Lomas J, Ahn H. A health opportunity cost threshold for cost-effectiveness analysis in the United States. *Ann Intern Med*. 2021;174(1):25–32. <https://doi.org/10.7326/M20-1392> PMID: [33136426](#)
25. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WEE, et al. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol*. 2016;11(1):39–51. <https://doi.org/10.1016/j.jtho.2015.09.009> PMID: [26762738](#)
26. Arias EX. United States life tables, 2020. National Vital Statistics Report. 2022;71(1).
27. Optum C. National Average Payment Table Update. 2023.
28. Agency for Healthcare Research and Quality. Mean Expenditure per Person by Age Groups, United States, 1996 to 2020.
29. National Comprehensive Cancer Network. Non-Small Cell Lung Cancer. 2023.
30. Centers for Medicare & Medicaid Services. 2021 ASP Drug Pricing Files. 2023.
31. Deppen SA, Blume JD, Kensinger CD, Morgan AM, Aldrich MC, Massion PP, et al. Accuracy of FDG-PET to diagnose lung cancer in areas with infectious lung disease: A meta-analysis. *JAMA*. 2014;312(12):1227–36. <https://doi.org/10.1001/jama.2014.11488> PMID: [25247519](#)
32. 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(10):910–9. <https://doi.org/10.1056/NEJMoa1214726> PMID: [24004118](#)
33. Farjah F, Monsell SE, Greenlee RT, Gould MK, Smith-Bindman R, Banegas MP, et al. Patient and nodule characteristics associated with a lung cancer diagnosis among individuals with incidentally detected lung nodules. *Chest*. 2023;163(3):719–30. <https://doi.org/10.1016/j.chest.2022.09.030> PMID: [36191633](#)
34. Tanner NT, Aggarwal J, Gould MK, Kearney P, Diette G, Vachani A, et al. Management of pulmonary nodules by community pulmonologists: A multicenter observational study. *Chest*. 2015;148(6):1405–14. <https://doi.org/10.1378/chest.15-0630> PMID: [26087071](#)
35. Aguinagalde B, Lizarbe IA, Rumbeo JC, Lorenzo M, Sanchez L, Sanchez L, et al. Is histological confirmation necessary to avoid futile resections? Comparative of 4 university hospitals. *Eur J Cardiothorac Surg*. 2024;65(4):eae136. <https://doi.org/10.1093/ejcts/ezae136> PMID: [38632054](#)
36. Altorki N, Wang X, Kozono D, Watt C, Landrenau R, Wigle D, et al. Lobar or sublobar resection for peripheral stage IA non-small-cell lung cancer. *N Engl J Med*. 2023;388(6):489–98. <https://doi.org/10.1056/NEJMoa2212083> PMID: [36780674](#)
37. Tanner NT, Porter A, Gould MK, Li X-J, Vachani A, Silvestri GA. Physician assessment of pretest probability of malignancy and adherence with guidelines for pulmonary nodule evaluation. *Chest*. 2017;152(2):263–70. <https://doi.org/10.1016/j.chest.2017.01.018> PMID: [28115167](#)