

Outcomes of Resected Lung Cancer Diagnosed Through Screening and Incidental Pulmonary Nodule Programs in a Mississippi Delta Cohort



Olawale Akinbobola, MPH,^a Wei Liao, PhD,^a Meredith A. Ray, PhD,^b Carrie Fehnel, BBA,^a Jordan Goss, MA,^a Talat Qureshi, BS,^a Andrea Saulsberry, DBA,^a Kourtney Dortch, BS,^a Matthew P. Smeltzer, PhD,^b Raymond U. Osarogiagbon, M.B.B.S, FACP^{a,*}

^aThoracic Oncology Research Group, Baptist Cancer Center, Memphis, Tennessee

^bSchool of Public Health, University of Memphis, Memphis, Tennessee

Received 24 April 2024; accepted 28 April 2024
Available online - 16 May 2024

ABSTRACT

Introduction: Early lung cancer detection programs improve surgical resection rates and survival but may skew toward more indolent cancers.

Methods: Hypothesizing that differences in stage-stratified survival indicate differences in biological aggressiveness and possible length-time and overdiagnosis bias, we assessed a cohort who had curative-intent resection, categorized by diagnostic pathways: screening, incidental pulmonary nodule program, and non-program based. Survival was analyzed using Kaplan-Meier plots, log-rank tests, and Cox regression, comparing aggregate and stage-stratified survival across cohorts with Tukey's method for multiple testing.

Results: Of 1588 patients, 111 patients (7%), 357 patients (22.5%), and 1120 patients (70.5%) were diagnosed through screening, pulmonary nodule, and non-program-based pathways; 0% versus 9% versus 6% were older than 80 years ($p = 0.0048$); 17%, 23%, and 24% had a Charlson Comorbidity score greater than or equal to 2 ($p = 0.0143$); 7%, 6%, and 9% had lepidic adenocarcinoma; 26%, 31%, and 34% had poorly or undifferentiated tumors ($p = 0.1544$); and 93%, 87%, and 77% had clinical stage I ($p < 0.0001$). Aggregate 5-year survival was 87%, 72%, and 65% ($p = 0.0009$), including 95%, 74%, and 74% for pathologic stage I. Adjusted pairwise comparisons showed similar survival in screening and nodule program cohorts ($p = 0.9905$). Nevertheless, differences were significant between screening and non-program-based cohorts ($p = 0.0007$, adjusted hazard ratio 0.33 [95% confidence interval: 0.18–0.6]) and between nodule and nonprogram cohorts (adjusted hazard ratio 0.77 [95% confidence interval: 0.61–0.99]). Stage I comparisons yielded $p =$

0.2256, 0.1131, and 0.911. In respective pathways, 0%, 2%, and 2% of patients with stage I disease who were older than 80 years had a Charlson score greater than or equal to 2 ($p = 0.3849$).

Conclusions: Neither length-time nor overdiagnosis bias was evident in NSCLC diagnosed through screening or incidental pulmonary nodule programs.

© 2024 Published by Elsevier Inc. on behalf of the International Association for the Study of Lung Cancer. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Length-time bias; Overdiagnosis bias; Early detection; Survival

Introduction

Early lung cancer detection, through screening and management of incidentally detected pulmonary

*Corresponding author.

Address for correspondence: Raymond U. Osarogiagbon, M.B.B.S, FACP, Thoracic Oncology Research Group, Multidisciplinary Thoracic Oncology Program, Baptist Cancer Center, 6141 Walnut Grove Road, Second Floor, Memphis, TN 38120. E-mail: rosarogi@bmhcc.org

Cite this article as: Akinbobola O, Liao W, Ray MA, et al. Outcomes of resected lung cancer diagnosed through screening and incidental pulmonary nodule programs in a Mississippi Delta cohort. *JTO Clin Res Rep* 2024;5:100684.

© 2024 Published by Elsevier Inc. on behalf of the International Association for the Study of Lung Cancer. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

ISSN: 2666-3643

<https://doi.org/10.1016/j.jtocrr.2024.100684>

nodules, increases the likelihood of curative-intent surgery and survival.¹⁻³ The risk profile of patients diagnosed through these two pathways differs, and the biological characteristics of the cancers detected may be dissimilar.^{4,5} Lung cancer screening decreases lung cancer-specific mortality but is more effective in diagnosing relatively indolent cancers, such as lepidic-predominant adenocarcinomas.^{1,6} This has raised questions about length-time bias—the preferential detection of slower-growing cancers due to their longer preclinical phase—and overdiagnosis bias—the identification of cancers that would not affect patient lifespan; these biases are particularly pertinent in screening-detected early-stage lung cancers, despite screening eligibility criteria being skewed toward patients at high risk.^{7,8} This undesired aspect of lung cancer screening has been variably estimated in clinical trial populations.⁷⁻¹²

Guideline-concordant management of incidentally detected pulmonary nodules expands the possibility of early lung cancer detection to a considerably more diverse population than currently reached by screening programs. Such populations include a predominance of persons ineligible for screening by current age and smoking history criteria.³⁻⁵ Nevertheless, the lung cancers diagnosed from lung nodules may also be more indolent than those diagnosed through conventional pathways of care. Furthermore, lung nodule management guidelines have no upper age limit of applicability, exacerbating concerns about overdiagnosis in persons with competing causes for mortality.¹³ To our knowledge, there are no established estimates of length-time and overdiagnosis bias among persons with lung cancer diagnosed through incidental pulmonary nodule programs. These biases, if present, would most likely be evident among recipients of curative-intent surgery for stage I NSCLC.¹⁴⁻¹⁶

We hypothesized that all things being equal, length-time bias would most likely manifest as a difference in program-based survival among patients with stage I disease, and overdiagnosis bias would be most likely in older patients with multiple co-morbidities. To estimate the likelihood of length-time and overdiagnosis bias in patients with lung cancer diagnosed through an incidental pulmonary nodule program, we compared the characteristics and survival of recipients of curative-intent surgical resection after diagnosis through three pathways: low-dose computed tomography screening, incidental pulmonary nodule program, and neither (non-program based).

Materials and Methods

Study Population

With the approval of the institutional review board of the Baptist Memorial Health Care Corporation (BMHCC),

including waiver of the informed consent requirement for this low-risk study, we evaluated the care and outcomes of all persons who had curative-intent lung cancer surgery within the healthcare system. BMHCC, a not-for-profit community-based healthcare system, serves a racially and socio-economically diverse population in more than 111 counties in Kentucky, Mississippi, Arkansas, Tennessee, Alabama, and Missouri. The service area population has some of the highest U.S. per capita lung cancer incidence and mortality rates.^{17,18}

Data Sets

We used data from two prospective observational institutional data sets, the Detecting Early Lung Cancer (DELUGE) in the Mississippi Delta and the Mid-South Quality of Surgical Resection (MS-QSR) databases, to identify patients who had curative-intent surgery for lung cancer after diagnosis through the three pathways. We have previously described each of these databases in detail.^{3,19} Briefly, the DELUGE database contains details on all patients who received a low-dose computed tomography scan for lung cancer screening, based on applicable eligibility criteria at the time of testing, and all patients managed in the Incidental Pulmonary Nodule Program, based on a radiologist's report of the presence of a potentially malignant lung lesion on radiologic studies performed for reasons other than known or suspected lung cancer.³⁻⁵ The MS-QSR data set is a population-level database of all patients who had curative-intent surgical resection for lung cancer in 14 hospitals in seven healthcare systems across five contiguous Hospital Referral Regions in Arkansas, Mississippi, and Tennessee, including six BMHCC hospitals.¹⁹ For this analysis, we restricted the MS-QSR data extract to the six BMHCC hospitals to align with the DELUGE cohort, which was entirely from BMHCC.

Patient Selection

Eligible patients underwent their first lung cancer resection in the MS-QSR from 2015 to 2022. We excluded those with previous lung cancer and considered only the first operation for patients with multiple resections, and recipients of neoadjuvant therapy were excluded to avoid confounding effects of prior treatments. By cross-linkage with the DELUGE database, we separated the eligible MS-QSR cohort into screened, nodule, and non-program-based cohorts, assuming those absent were diagnosed without the benefit of program-based early detection.

Bias Analysis

The survival benefit of early lung cancer detection derives from redistribution to earlier stage, but all things being equal, stage-stratified survival should be

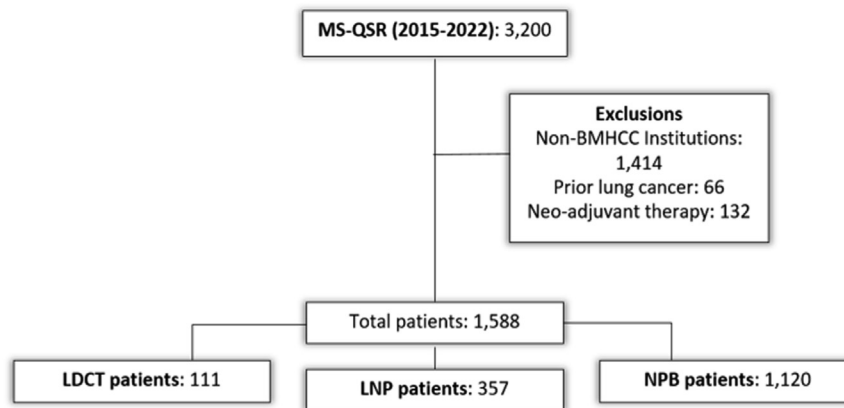


Figure 1. Cohort selection diagram—patient selection and categorization into program-based pathways of lung cancer diagnosis. BMHCC, Baptist Memorial Health Care Corporation; LDCT, low-dose computed tomography; LNP, lung nodule program; MS-QSR, Mid-South Quality of Surgical Resection; NPB, non-program-based.

equivalent irrespective of pathway of diagnosis. We therefore hypothesized that excess stage-stratified survival in patients with early-stage disease would indicate length-time bias. We further hypothesized that overdiagnosis bias would be prevalent in the extremely older patients with significant co-morbidities and competing causes of mortality. To estimate overdiagnosis bias, we compared the proportion of patients with lung cancer older than 80 years at the time of diagnosis, who also had a Charlson Comorbidity score greater than or equal to 2.

Outcomes

The primary outcome was overall survival from the date of surgery to date of death. A secondary outcome was 30-, 60-, 90-, and 120-day postoperative mortality.

Statistical Analysis Plan

We summarized cohort characteristics and secondary outcomes as frequencies, percentages, means, medians, and interquartile ranges (IQR). We used the chi-square and Fisher's exact tests to compare characteristics among screened, nodule, and non-program-based cohorts. We compared medians between programs using the Wilcoxon-Mann-Whitney test. We used the Kaplan-Meier method to visualize the differences in survival among the three pathways and compared survival using the log-rank test. We performed pairwise comparisons of survival curves to assess the differences in survival between programs, adjusting for multiple comparisons using Tukey's method. We used Cox proportional regression to estimate crude and adjusted hazard ratios with 95% confidence intervals (CIs) between programs. $p < 0.05$ was considered statistically significant, and all analyses were performed using SAS Version 9.4 (Cary, North Carolina).

Results

Cohort Characteristics

From January 2015 to December 2022, there were 3200 curative-intent resections in the MS-QSR database, from which we excluded 1414 resections performed outside the six BMHCC hospitals, 66 resections with previous lung cancer, and 132 with neoadjuvant therapy (Fig. 1). The analysis cohort comprised 1588 patients, including 111 patients (7%) in the screened cohort, 357 patients (22.5%) in the nodule cohort, and 1120 patients (70.5%) in the non-program-based cohort. The screened, nodule, and non-program-based cohorts had a median age of 69 years (IQR: 64–73), 70 years (IQR: 64–76), and 69 years (IQR: 63–75), respectively. The screened cohort had no individuals older than 80 years, whereas 9% and 6% of the nodule and non-program-based cohorts were older than 80 years. In addition, the screened cohort had a higher proportion of persons of White race (84%) than did the nodule and non-program-based cohorts (77%) (Table 1). There was no difference in the distribution of sex or insurance among cohorts.

In total, 60% of patients who were screened actively smoked, compared with 38% of the nodule and 40% of the non-program-based cohorts. By eligibility criteria, all patients screened had a history of smoking, whereas 18% of the nodule and 13% of the non-program-based cohorts had never smoked ($p < 0.0001$). The distribution of Charlson Comorbidity scores differed among cohorts: 17% versus 23% versus 24% had a score greater than or equal to 2 ($p = 0.0143$).

Lung Cancer Evaluation

The radiologic lesion was smallest in the screened cohort and largest in the non-program-based cohort:

Table 1. Demographic and Clinical Characteristics of Patients in the Pathways of Care

Patient-Level Variable	Lung Cancer Screening	Lung Nodule Program	Non-Program Based	p-Value
	n = 111	n = 357	n = 1120	
Age, n (%)				0.0048
<49	-	6 (2)	33 (3)	
50-64	29 (26)	84 (24)	296 (26)	
65-80	82 (74)	234 (66)	722 (64)	
>80	-	33 (9)	69 (6)	
Mean (SD)	68.4 (5.7)	69.7 (9.1)	68.4 (9)	0.0205
Median (Q1, Q3)	69 (64, 73)	70 (64, 76)	69 (63, 75)	
Race, n (%)				0.0905
White	93 (84)	274 (77)	863 (77)	
Black	18 (16)	71 (20)	243 (22)	
Other	-	10 (3)	12 (1)	
Not reported	-	2 (1)	2 (0)	
Sex, n (%)				0.2347
Male	51 (46)	157 (44)	549 (49)	
Female	60 (54)	200 (56)	571 (51)	
Insurance, n (%)				0.7819
Medicare	52 (47)	171 (48)	541 (48)	
Medicaid	13 (12)	45 (13)	168 (15)	
Commercial	45 (41)	135 (38)	397 (35)	
Self-Insured/None	1 (1)	6 (2)	14 (1)	
Smoking Status				<0.0001
Active	67 (60)	134 (38)	453 (40)	
Former	44 (40)	158 (44)	516 (46)	
Never	-	65 (18)	149 (13)	
Not Reported	-	-	2 (0)	
Charlson Comorbidity Score, n (%)				0.0143
0	19 (17)	102 (29)	301 (27)	
1	73 (66)	174 (49)	548 (49)	
≥2	19 (17)	81 (23)	271 (24)	
Co-morbidities				0.292
Mean (SD)	2.3 (1)	2.2 (1.2)	2.2 (1.2)	
Median (Q1, Q3)	2 (2, 3)	2 (1, 3)	2 (1, 3)	

SD, standard deviation; Q, quartile.

68%, 57%, and 47% were 0 to 2 cm, and the median tumor size was 1.6 cm (IQR: 1.1–2.3) versus 1.9 cm (IQR: 1.4–2.8) versus 2.2 cm (IQR: 1.4–3.1), respectively, $p < 0.0001$ (Table 2). Positron-emission tomographic-computer tomographic scans were less typically performed in the early detection cohorts—86% versus 87% versus 94%, $p < 0.0001$; there was no difference in the use of invasive staging among cohorts—13% versus 18% versus 17% ($p = 0.3544$).

Lung Cancer Characteristics

Although adenocarcinoma was least common in patients screened for lung cancer, 50% versus 61% and 60% in the nodule and non-program-based cohorts, the distribution of lepidic adenocarcinoma was similar—7% versus 6% versus 9%. Squamous cell carcinoma was most common in the screened cohort—44% versus 25% versus 28% (Table 2). Although 8%, 16%, and 14% of lung cancers were well

differentiated, and 26%, 31%, and 34% were poorly differentiated or undifferentiated, this was not significantly different ($p = 0.1544$). The tumor grade was not known in 10% to 11% of patients.

Clinical stage was most skewed toward early stage in patients screened: 85%, 77%, and 64% of patients had T1 tumors, and 84%, 76%, and 62% had stage IA disease ($p < 0.0001$ for clinical T-category comparison). Although there was no significant difference in the N category distribution, with 96%, 97%, and 94% clinical N0 ($p = 0.404$), aggregate pathologic stage also differed significantly, with 84%, 76%, and 63% pathologic stage I, 8%, 16%, and 21% stage II, and 8%, 8%, and 15% stage III ($p < 0.0001$). This difference was attributable to differences in the pathologic T-category with 68% versus 59% versus 48% with pathologic T-in situ to T1 and no significant difference in pathologic N-category—86% versus 85% versus 79% were pathologic N0 ($p = 0.1515$).

Table 2. Lung Cancer, Treatment Characteristics, and Postoperative Mortality Comparison Among Cohorts

Disease, Treatment Variables, and Mortality	Screened N = 111	Nodule N = 357	Non-Program Based N = 1120	p-Value
Disease and Treatment Characteristics				
Histology, n (%)				0.0154
Adenocarcinoma	55 (50)	219 (61)	677 (60)	
Adeno-Lepidic	8 (7)	22 (6)	97 (9)	
Squamous	49 (44)	89 (25)	309 (28)	
Adenosquamous	1 (1)	8 (2)	25 (2)	
Large Cell	2 (2)	9 (3)	18 (2)	
Other	4 (4)	32 (9)	91 (8)	
Tumor Grade, n (%)				0.1544
Well differentiated	9 (8)	58 (16)	157 (14)	
Moderately differentiated	61 (55)	151 (42)	471 (42)	
Poorly differentiated	28 (25)	108 (30)	360 (32)	
Undifferentiated	1 (1)	4 (1)	24 (2)	
Not Reported/Cannot Be Assessed	12 (11)	36 (10)	108 (10)	
Tumor Size (cm), n (%)				<0.0001
0-2	76 (68)	202 (57)	530 (47)	
>2-3	19 (17)	87 (24)	290 (26)	
>3-5	12 (11)	62 (17)	210 (19)	
>5-7	4 (4)	6 (2)	72 (6)	
≥7	-	-	18 (2)	
Preradiologic Tumor Size (cm)				<0.0001
Mean (SD)	1.9 (1.1)	2.2 (1.1)	2.5 (1.6)	
Median (Q1, Q3)	1.6 (1.1, 2.3)	1.9 (1.4, 2.8)	2.2 (1.4, 3.1)	
Postpathologic Tumor Size (cm)				<0.0001
Mean (SD)	2.1 (1.3)	2.4 (1.3)	2.9 (1.9)	
Median (Q1, Q3)	1.8 (1.3, 2.5)	2 (1.5, 2.8)	2.3 (1.6, 3.5)	
PET/CT, n (%)				<0.0001
No	15 (14)	45 (13)	68 (6)	
Yes	96 (86)	312 (87)	1052 (94)	
Invasive Staging, n (%)				0.3544
No	97 (87)	291 (82)	930 (83)	
Yes	14 (13)	66 (18)	190 (17)	
Clinical Stage, n (%)				<0.0001
Stage 0	-	-	3 (0)	
Stage IA1-IA3	93 (84)	270 (76)	690 (62)	
Stage IB	10 (9)	41 (11)	163 (15)	
Stage IIA-IIB	6 (5)	36 (10)	146 (13)	
Stage IIIA-IIIB	2 (2)	6 (2)	64 (6)	
Stage IV	-	2 (1)	22 (2)	
Stage Not Known	-	2 (1)	32 (3)	
Clinical T Stage, n (%)				<0.0001
T0, Tx, Tis	-	-	5 (0%)	
T1	94 (85)	275 (77)	718 (64)	
T2	12 (11)	62 (17)	227 (20)	
T3	5 (5)	17 (5)	101 (9)	
T4	-	1 (0)	39 (3)	
Insufficient records	-	2 (1)	30 (3)	
Clinical N Stage, n (%)				0.404
N0	107 (96)	345 (97)	1051 (94)	
N1	4 (4)	7 (2)	29 (3)	
N2	-	3 (1)	25 (2)	
N3	-	-	2 (0)	
Insufficient records	-	2 (1)	13 (1)	

(continued)

Table 2. Continued

Disease, Treatment Variables, and Mortality	Screened N = 111	Nodule N = 357	Non-Program Based N = 1120	p-Value
Clinical M Stage, n (%)				0.281
M0	111 (100)	353 (99)	1086 (97)	
M1a	-	1 (0)	3 (0)	
M1b	-	1 (0)	19 (2)	
Insufficient records	-	2 (1)	12 (1)	
Pathologic Stage, n (%)				<0.0001
Stage IA1-IA3	68 (61)	189 (53)	480 (43)	
Stage IB	25 (23)	83 (23)	225 (20)	
Stage IIA-IIB	9 (8)	56 (16)	236 (21)	
Stage IIIA-IIIB	9 (8)	27 (8)	167 (15)	
Stage IV	-	2 (1)	12 (1)	
Pathologic T Stage, n (%)				<0.0001
T0, Tx, Tis	1 (1)	3 (1)	4 (0)	
T1	74 (67)	206 (58)	536 (48)	
T2	28 (25)	110 (31)	357 (32)	
T3	7 (6)	35 (10)	162 (14)	
T4	1 (1)	3 (1)	61 (5)	
Pathologic N Stage, n (%)				0.1515
Nx	2 (2)	5 (1)	33 (3)	
N0	95 (86)	304 (85)	886 (79)	
N1	7 (6)	29 (8)	109 (10)	
N2	7 (6)	19 (5)	92 (8)	
Pathologic M Stage, n (%)				0.6132
M0	110 (99)	354 (99)	1108 (99)	
M1a	1 (1)	2 (1)	4 (0)	
M1b	-	1 (0)	8 (1)	
Final Surgical Technique, n (%)				<0.0001
Open	35 (32)	84 (24)	364 (33)	
Robotically-assisted thoracoscopic surgery	60 (54)	241 (68)	433 (39)	
Video-assisted thoracoscopic surgery	16 (14)	32 (9)	322 (29)	
Not Reported	-	-	1 (0)	
Extent of Resection, n (%)				0.0032
Pneumonectomy	2 (2)	2 (1)	19 (2)	
Bilobectomy	1 (1)	4 (1)	28 (3)	
Lobectomy	92 (83)	317 (89)	959 (86)	
Segmentectomy	6 (5)	24 (7)	37 (3)	
Wedge	10 (9)	10 (3)	77 (7)	
Adjuvant Therapy	26 (23)	81 (23)	307 (27)	0.1682
Postoperative Mortality				
30-D Mortality, n (%)				0.2914
No	111 (100)	349 (98)	1100 (98)	
Yes	0 (0)	8 (2)	20 (2)	
60-D Mortality, n (%)				0.1667
No	111 (100)	347 (97)	1085 (97)	
Yes	0 (0)	10 (3)	35 (3)	
90-D Mortality, n (%)				0.0339
No	111 (100)	345 (97)	1070 (96)	
Yes	0 (0)	12 (3)	50 (4)	
120-D Mortality, n (%)				0.0689
No	110 (99)	344 (96)	1061 (95)	
Yes	1 (1)	13 (4)	59 (5)	

CT, computed tomography; PET, positron emission tomography; Q, quartile.

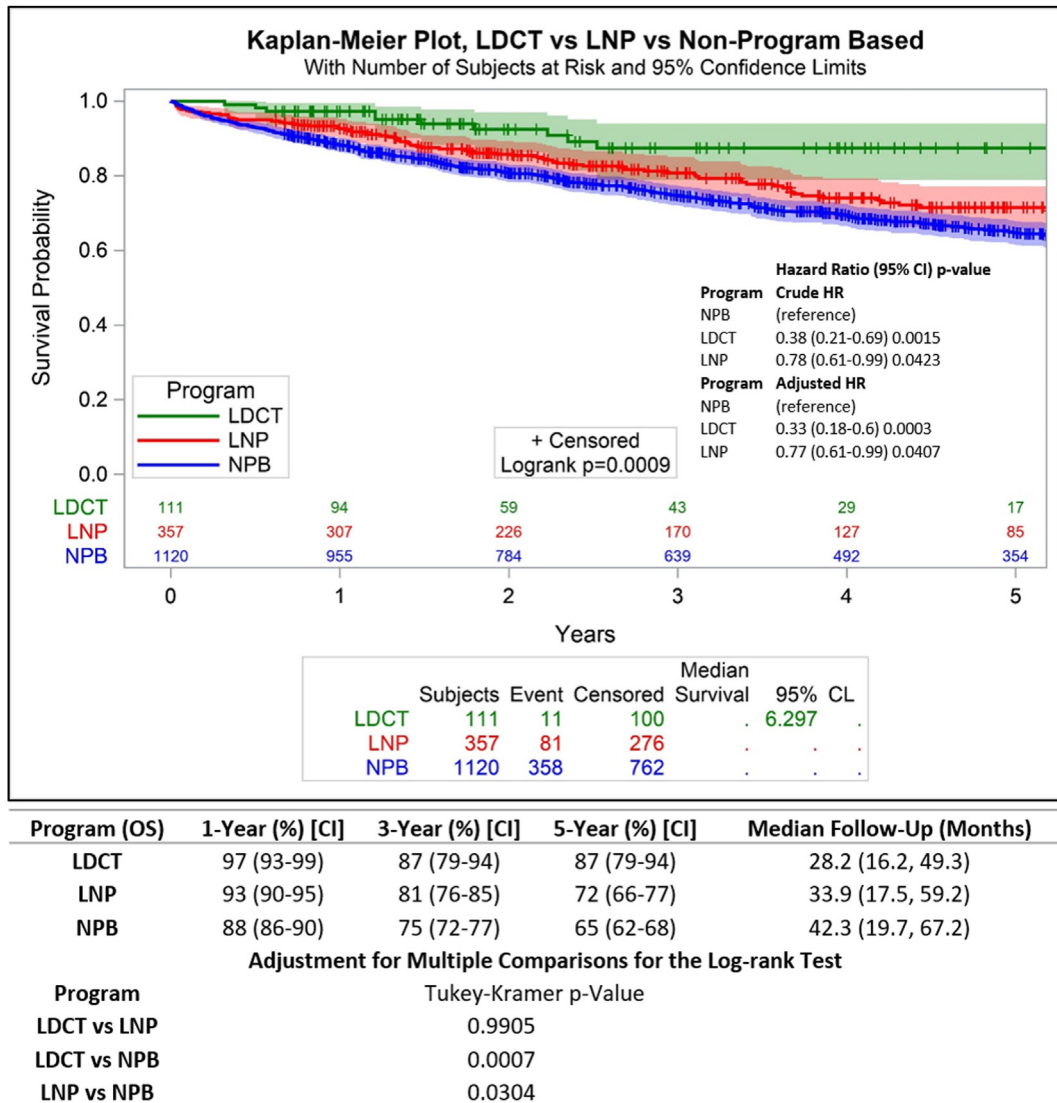


Figure 2. Overall survival comparisons and hazard ratio of patients who had curative-intent surgical resection according to pathway of diagnosis. CI, confidence interval; CL, confidence limit; HR, hazard ratio; LDCT, low-dose computed tomography; LNP, lung nodule program; NPB, non-program-based.

Treatment

Surgical resection was minimally invasive in 68% versus 77% versus 68% and open in 32% versus 24% versus 33% ($p < 0.0001$). The extent of resection differed significantly among cohorts, with bilobectomy or pneumonectomy in 3% versus 2% versus 5% and sublobar resection in 14% versus 10% versus 10% ($p = 0.0032$). Adjuvant therapy did not significantly differ among cohorts—23% versus 23% versus 27% ($p = 0.1682$; Table 2).

Outcomes

Surgery for lung cancer was very safe in the screened cohort, with no postoperative deaths at 90 days and one death within 120 days

(Table 2). Although not significantly different across the three cohorts, 30- and 60-day postoperative mortality rates were 2% and 3% in both the nodule and non-program-based cohorts. Nevertheless, postoperative mortality rates differed significantly at 90 days, 0 versus 3% versus 4% ($p = 0.0339$), and at 120 days, 1% versus 4% versus 5% ($p = 0.0689$).

After a median of 28-, 34-, and 42-months follow-up in the screened, nodule, and non-program-based cohorts, the median survival was not reached in any of the cohorts. Aggregate 5-year overall survival was 87% (95% CI: 79–94) versus 72% (95% CI: 66–77) versus 65% (95% CI: 62–68) in the respective cohorts (log-rank $p = 0.0009$; Fig. 2). In the pair-wise comparison of survival adjusting for multiple testing, there was no

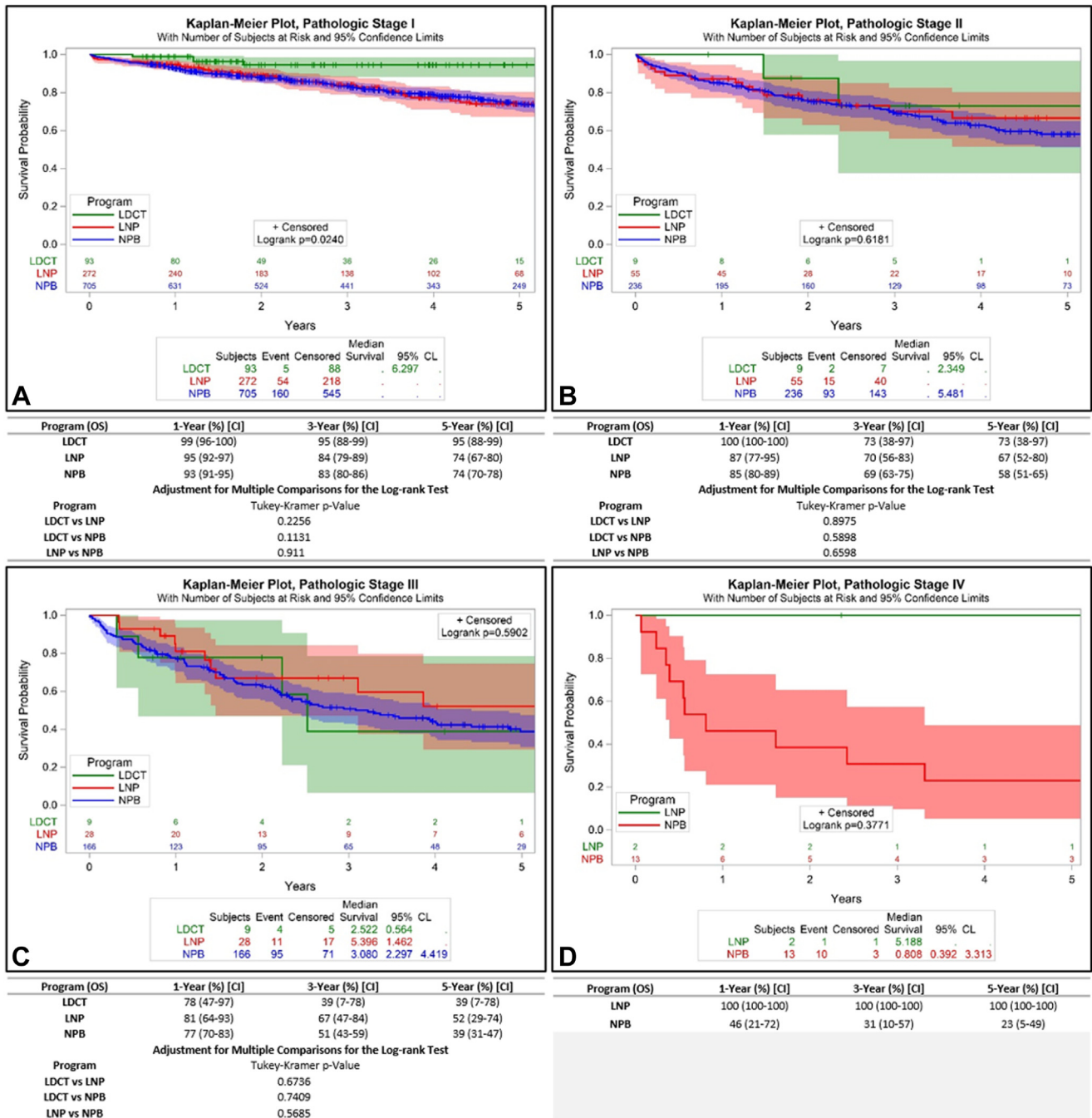


Figure 3. Pathologic stage-stratified survival of patients who undergo curative-intent surgical resection according to pathway of diagnosis. (A) Pathologic stage I; (B) Pathologic stage II; (C) Pathologic stage III; (D) Pathologic stage IV. CI, confidence interval; CL, confidence limit; LDCT, low-dose computed tomography; LNP, lung nodule program; NPB, non-program-based.

significant difference between the screened and nodule cohorts ($p = 0.9905$), but survival was worse in the non-program-based cohort than in the screened ($p = 0.0007$) and nodule ($p = 0.0304$) cohorts.

In the screened and nodule cohorts, the crude hazard ratio for death compared with the non-program-based cohort was 0.38 (95% CI: 0.21–0.69) and 0.78 (95% CI: 0.61–0.99), respectively ($p < 0.05$

unless noted). The results were consistent after adjusting for age, sex, race, smoking history, Charlson Comorbidity score, extent of resection, histologic type, and tumor grade, with the screened cohort showing a 67% reduction in the hazard of death 0.33 (95% CI: 0.18–0.6) and the nodule cohort showing a 23% reduction in the hazard of death 0.77 (95% CI: 0.61–0.99; Fig. 2).

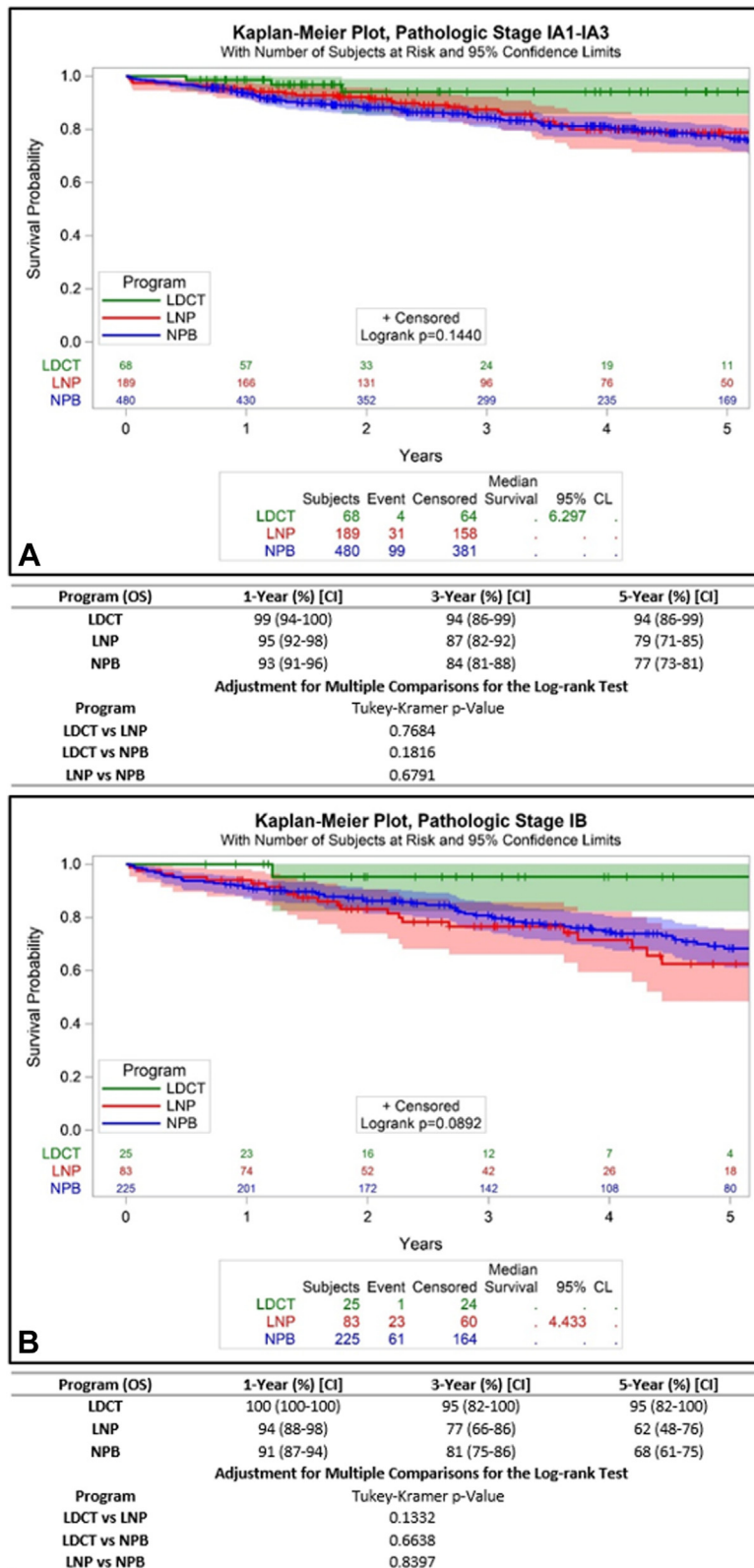


Figure 4. Pathologic stage-stratified survival of patients who undergo curative-intent surgical resection according to pathway of diagnosis. (A) Pathologic stage IA; (B) Pathologic stage IB. CL, confidence limit; LDCT, low-dose computed tomography; LNP, lung nodule program; NPB, non-program-based.

Stage Distribution and Stage-Stratified Survival

We attempted to determine whether aggregate survival differences were attributable mostly to stage distribution differences or within-stage survival differences. In the screened cohort, 84%, 8%, 8%, and 0% had pathologic stage I, II, III, and IV, respectively, compared with 76%, 16%, 8%, and 1% of the nodule cohort, and 63%, 21%, 15%, and 1% of the non-program-based cohort. To test the hypothesis regarding length-time bias, we compared pathologic stage-stratified survival among the three cohorts. The 5-year survival rates of patients with pathologic stage I was 95% for the screened cohort, versus 74% and 74% ($p = 0.0240$; Fig. 3A), suggesting the possibility of length-time bias in the screened but not the nodule cohort. Survival rates were similar in patients with pathologic stage II, III, and IV. Specifically, stage II survival rates were 73%, 67%, and 58%, respectively ($p = 0.6181$); stage III rates were 39%, 52%, and 39% ($p = 0.5902$); and stage IV rates were 100% and 23% for the screened and non-program-based cohorts ($p = 0.3771$) (Fig. 3B–D). To further evaluate the possibility of length-time bias, we compared the survival of the stage I cohorts stratified into IA and IB, of which the respective proportions were 61% and 23% in the screened cohort, 53% and 23% in the nodule cohort, and 43% and 20% in the non-program-based cohort. The 5-year survival rates for IA were 94% versus 79% v 77%, ($p = 0.1440$), and survival of IB was 95% versus 62% versus 68% ($p = 0.0892$) (Fig. 4A and B).

Comparative Aggregation of Older Age and Multiple Co-morbidities

There were no patients in the screened cohort older than 80 years, compared with 9% and 6% in the nodule and non-program-based cohorts, respectively; of these, 2% in both cohorts also had a Charlson Comorbidity score greater than or equal to 2 (Table 2).

Discussion

In this prospective curative-intent lung cancer surgical resection cohort segregated according to pathway of lung cancer detection, patients diagnosed through an incidental pulmonary nodule program had demographic and clinical characteristics that were more similar to those of patients conventionally diagnosed without the benefit of early detection and differed somewhat from patients screened in several respects, including in the distribution of age, race, smoking history, Charlson Comorbidity scores, and tumor histologic type. The proportion of patients with lepidic adenocarcinoma and the distribution of tumor grade was similar across the three cohorts. The distribution of clinical and pathologic stage of the nodule program

cohort was intermediate between the screened and non-program-based cohorts.

Surgery was relatively safe across the three cohorts, with 90-day mortality rates ranging from 0% in the screened cohort to 4% in the non-program-based cohort. The 120-day mortality rate of 1% in this screened cohort from a community healthcare system is reassuring, given the recommended restriction of eligibility to persons who are asymptomatic and in relatively good health. Patients with incidentally detected pulmonary nodules typically had a clinical need for a radiologic study from which the malignant lesion was eventually detected. In theory, such preexisting active health problems can introduce delays in lung cancer care and may represent competing causes of morbidity and mortality.

Even though all patients had surgical resection, there were differences in survival according to the pathway of cancer detection, with those detected through the early detection programs having better survival than did the patients in the non-program-based pathway. The primary reason for this seems to be the relatively greater skew toward earlier clinical and pathologic stage in the early detection cohorts. Interestingly, in pairwise analysis adjusting for multiple testing, there was no difference in survival between the screened and pulmonary nodule cohorts. This differs from analyses including all patients, irrespective of treatment modality (and therefore including the full spectrum of stage and physical condition), in which the screened cohorts had better survival than did the pulmonary nodule program cohort.^{3,20}

To inform the ongoing debate about how much of the survival benefit of lung cancer screening is attributable to lead-time bias—the apparent increase in survival time due to earlier detection rather than a true clinical benefit—in clinical trials and in implementation reports of lung cancer screening in non-clinical trial populations, we sought to uncover evidence of similar bias in the screened and nodule program cohorts in our Mississippi Delta population.^{6–16,21,22} Hypothesizing that overdiagnosis bias would be most evident in a stage I surgical resection NSCLC cohort, we compared the stage-stratified survival of patients diagnosed through the two early detection programs with that of patients diagnosed without the benefit of a structured early detection program. The screened cohort with stage I showed a trend toward improved survival, as observed in the Kaplan-Meier plot. Upon further evaluation, this seemed to be mostly attributable to a higher proportion of stage IA in the screened group. The stage-stratified survival of the nodule program cohort overlapped that of the patients conventionally diagnosed, refuting the length-time bias hypothesis.

Lung nodule management guidelines, such as those of the Fleischner Society, have no upper age limit, raising the possibility that the incidental pulmonary nodule programs might have an excess of older patients who are frail and undergo surgical resection, with limited long-term benefit.^{5,13} We estimated this possibility by comparing the prevalence of patients with the combination of age advanced beyond screening eligibility (80 years) and significant co-morbidities (a Charlson score of ≥ 2) in the nodule and non-program-based cohorts. We found no meaningful difference between the cohorts in this crude estimate of overdiagnosis bias.

These findings augur well for the potential population-level impact of structured early lung cancer detection programs, especially in relatively high-risk populations such as in the Mississippi Delta. The greater the degree of length-time and overdiagnosis bias associated with early detection programs, the less cost-effective would be such programs.^{6,10,12} Our findings suggest that the lung cancers diagnosed through the nodule program were as biologically aggressive as those diagnosed through conventional means, providing some reassurance that the increased volume of cases in the nodule program is likely to have the salubrious public health effect of reducing population-level lung cancer mortality. The main explanation for improved survival in the early detection programs seems to be redistribution toward earlier stage, not diagnosis of biologically indolent cancers.

Limitations

Although prospectively collected, the DELUGE and MS-QSR databases were not constructed to test the hypotheses in the current analysis, which is therefore susceptible to biases inherent in retrospective analyses. Because the data set lacks information on cause-specific mortality and lung cancer recurrence, we cannot account for competing causes of mortality, which may yet explain overall survival differences among the cohorts and mask differences between the nodule and non-program-based cohorts. The differing median follow-up times across cohorts may introduce follow-up bias, potentially influencing survival outcome comparisons, and the smaller size of the screening cohort than of the non-program-based cohort may limit the study's power to detect true differences or similarities in survival outcomes. Data limitations will not permit a direct assessment of lead-time or overdiagnosis bias. There is a potential bias to the null from misclassification—some patients in the non-program-based cohort may have been referred for surgery within our program after diagnosis through screening or nodule management outside our program. The presence of such patients could mask real differences in stage-stratified survival among the cohorts.

Nevertheless, their presence would also narrow the apparent aggregate survival differences among pathways of diagnosis. Finally, we are not currently able to report on population-level cause-specific mortality differences, the ultimate accepted standard for measuring the impact of early cancer detection programs.

These limitations notwithstanding, we provide reasonable reassurance that community-level implementation of lung cancer screening and programs to foster guideline-concordant management of pulmonary nodules are safe and effective in patients who undergo curative-intent surgical resection. Redistribution toward earlier stage seems to be the main explanation for the survival benefit. Length-time and overdiagnosis bias, if present, seemed minimal in this cohort, providing reassurance that rigorous implementation of such programs will have the desired impact of reducing population-level lung cancer mortality. In ongoing work, we plan to closely examine the biological characteristics of the lung cancers detected through the different pathways as a means of understanding their underlying causal mechanisms given the striking differences in risk factor profile among cohorts.

CRedit Authorship Contribution Statement

Olawale Akinbobola: Data curation, Formal analysis, Methods, Software, Validation, Writing - original draft, review, and editing.

Wei Liao: Resources, Writing - review and editing.

Meredith A. Ray: Formal analysis, Methods, Validation, Writing - review and editing.

Carrie Fehnel: Data curation, Supervision, Project administration, Writing - review and editing.

Jordan Goss: Data curation, Writing - review and editing.

Talat Qureshi: Data curation, Writing - review and editing.

Andrea Saulsberry: Data curation, Writing - review and editing.

Kourtney Dortch: Data curation, Writing - review and editing.

Matthew Smeltzer: Formal analysis, Methods, Validation, Writing - review and editing.

Raymond U. Osarogiagbon: Conceptualization, Funding acquisition, Investigation, Visualization, Methods, Resources, Supervision, Validation, Writing - original draft, review, and editing.

Disclosure

Dr. Matthew P. Smeltzer is a paid research consultant for the Association of Community Cancer Centers. Dr. Raymond U. Osarogiagbon holds patents for surgical

specimen collection kit and stocks in Pfizer, Gilead Sciences, and Eli Lilly; is a paid research consultant for the American Cancer Society, the Association of Community Cancer Centers, Genentech/Roche, Biondix, Lungevity Foundation, National Cancer Institute, Tryplich Healthcare Partners, and AstraZeneca; and is founder of Oncobox Device, Inc. The remaining authors declare no conflict of interest.

Acknowledgments

This project was supported by grants from the Baptist Memorial Health Care Foundation for initial funding support (Grant #15BD03).

References

1. National Lung Screening Trial Research Team, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365:395-409.
2. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med*. 2020;382:503-513.
3. Osarogiagbon RU, Liao W, Faris NR, et al. Lung cancer diagnosed through screening, lung nodule, and neither program: a prospective observational study of the detecting early lung cancer (DELUGE) in the Mississippi Delta cohort. *J Clin Oncol*. 2022;40:2094-2105.
4. Osarogiagbon RU, Liao W, Faris NR, et al. Evaluation of lung cancer risk among persons undergoing screening or guideline-concordant monitoring of lung nodules in the Mississippi Delta. *JAMA Netw Open*. 2023;6:e230787.
5. Liao W, Fehnel C, Goss J, et al. Incidentally detected lung cancer in persons too young or too old for lung cancer screening in a Mississippi Delta cohort. *J Thorac Oncol*. 2024;19:589-600.
6. Black WC, Gareen IF, Soneji SS, et al. Cost-effectiveness of CT screening in the National Lung Screening Trial. *N Engl J Med*. 2014;371:1793-1802.
7. Patz EF Jr, Pinsky P, Gatsonis C, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. *JAMA Intern Med*. 2014;174:269.
8. National Lung Screening Trial Research Team. Lung cancer incidence and mortality with extended follow-up in the National lung screening trial. *J Thorac Oncol*. 2019;14:1732-1742.
9. Grannis FW Jr. Are we underestimating efficacy and cost-efficacy of population lung cancer computed tomography screening? *JTO Clin Res Rep*. 2022;3:100383.
10. Ten Haaf K, de Koning HJ. Overdiagnosis in lung cancer screening: why modelling is essential. *J Epidemiol Community Health*. 2015;69:1035-1039.
11. Durham DD, Crowell JM, Marcus PM. Do competing causes of mortality contribute to overdiagnosis in lung cancer screening? *Lung Cancer*. 2021;153:21-24.
12. Ten Haaf K, Tammemägi MC, Bondy SJ, et al. Performance and cost-effectiveness of computed tomography lung cancer screening scenarios in a population-based setting: a microsimulation modeling analysis in Ontario, Canada. *PLoS Med*. 2017;14:e1002225.
13. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology*. 2017;284:228-243.
14. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst*. 2010;102:605-613.
15. Baker SG, Prorok PC, Kramer BS. Lead time and overdiagnosis. *J Natl Cancer Inst*. 2014;106:dju346.
16. Dunn BK, Woloshin S, Xie H, Kramer BS. Cancer overdiagnosis: a challenge in the era of screening. *J Natl Cancer Cent*. 2022;2:235-242.
17. Mokdad AH, Dwyer-Lindgren L, Fitzmaurice C, et al. Trends and patterns of disparities in cancer mortality among US counties, 1980-2014. *JAMA*. 2017;317:388-406.
18. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. 2024;74:12-49.
19. Osarogiagbon RU, Smeltzer MP, Faris NR, et al. Outcomes after use of a lymph node collection kit for lung cancer surgery: a pragmatic, population-based, multi-institutional, staggered implementation study. *J Thorac Oncol*. 2021;16:630-642.
20. Liao W, Ray M, Fehnel C, et al. Program-based lung cancer care: a prospective observational tumor registry linkage study. *JTO Clin Res Rep*. 2023;5:100629.
21. Yang SC, Wang JD, Wang SY. Considering lead-time bias in evaluating the effectiveness of lung cancer screening with real-world data. *Sci Rep*. 2021;11:12180.
22. Gao W, Wen CP, Wu A, Welch HG. Association of computed tomographic screening promotion with lung cancer overdiagnosis among Asian women. *JAMA Intern Med*. 2022;182:283.