



Published in final edited form as:

*Semin Thorac Cardiovasc Surg.* 2021 ; 33(1): 206–216. doi:10.1053/j.semtcvs.2020.08.006.

## How Effective Is Neoadjuvant Therapy Followed by Surgery for Pathologic Single-Station N2 Non-Small Cell Lung Cancer?

Hari B. Keshava, MD, MS<sup>1,\*</sup>, Kay See Tan, PhD<sup>2,\*</sup>, Joseph Dycoco, BS<sup>1</sup>, Jennifer Livschitz, BS<sup>1</sup>, Matthew J. Bott, MD<sup>1</sup>, James Huang, MD, MS<sup>1</sup>, Valerie W. Rusch, MD<sup>1</sup>, James M. Isbell, MD<sup>1</sup>, Daniela Molena, MD<sup>1</sup>, Manjit S. Bains, MD<sup>1</sup>, David R. Jones, MD<sup>1</sup>, Gaetano Rocco, MD<sup>1</sup>

<sup>1</sup>Thoracic Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY

<sup>2</sup>Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY

### Abstract

**Objective:** The optimal treatment strategy for pathologic single-station N2 (pN2a1) non-small cell lung cancer (NSCLC)—surgery first followed by adjuvant treatment (SF) or neoadjuvant therapy followed by surgery (NS)—remains unclear. We compared disease-free survival (DFS) and overall survival (OS) after NS versus SF for pN2a1 NSCLC.

**Methods:** We retrospectively identified patients with pN2a1 NSCLC resected between 2000 and 2018. Patients in the SF group had cN0 disease and were treated with surgery before adjuvant chemotherapy; patients in the NS group had known preoperative nodal disease, cN2 disease, and were treated with neoadjuvant therapy before surgery. The matching-weights procedure was applied to generate a cohort with similar characteristics between groups. DFS and OS were calculated using the Kaplan-Meier approach and compared between groups using weighted log-rank test and Cox proportional hazards models.

**Results:** We identified 227 patients with pN2a1 disease: 121 treated with SF and 106 with NS. After the matching-weights procedure, 5- and 10-year DFS were 45% and 27% for SF versus 26% and 21% for NS (log-rank  $p=0.056$ ; hazard ratio [HR], 1.61; 95% confidence interval [CI],

---

**Corresponding Author:** Gaetano Rocco, MD, FRCSEd, Thoracic Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, Suite C-861, New York, NY 10065. Phone: 212-639-3478; roccog@mskcc.org.

\*H.B.K. and K.S.T. share first authorship.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Read at the American Association for Thoracic Surgery 99th Annual Meeting, Toronto, Canada, May 4-7, 2019

**IRB approval:** IRB #18-391; approved 9/7/18.

Webcast

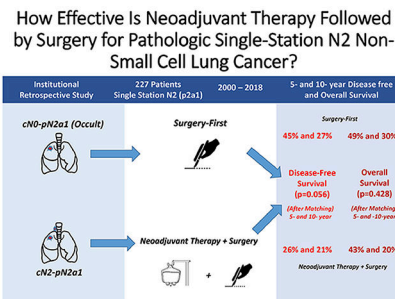
You can watch a Webcast of this MEETING NAME meeting presentation by going to: webcast address to be placed here:

[https://aats.blob.core.windows.net/media/19%20AM/Monday\\_May6/201DF/201DF/S84%20-%20Locally%20advanced%20lung%20cancer/S84\\_5\\_webcast\\_025849596.mp4](https://aats.blob.core.windows.net/media/19%20AM/Monday_May6/201DF/201DF/S84%20-%20Locally%20advanced%20lung%20cancer/S84_5_webcast_025849596.mp4)

0.98-2.65); 5- and 10-year OS were 49% and 30% for SF versus 43% and 20% for NS (log-rank  $p=0.428$ ; HR, 1.24; 95% CI, 0.67-2.28).

**Conclusions:** SF and NS for pN2a1 NSCLC resulted in similar survival. A study comparing SF for known presectional pN2a1 with occult pN2a1 disease could be a next step. Further investigation of SF for known N2a1 versus occult pN2a1 disease could power a clinical trial focused on N2a NSCLC.

## Graphical Abstract



## Keywords

non-small cell lung cancer; pathologic single-station N2; adjuvant treatment; neoadjuvant treatment; survival

## INTRODUCTION

In an analysis of 26,000 patients performed for the revised proposal on nodal (N) status descriptors in the 8<sup>th</sup> edition of the tumor-node-metastasis (TNM) classification for non-small cell lung cancer (NSCLC), 5-year overall survival (OS) for pathologic N2 (pN2) disease was 38%, with a median survival time (MST) of 39 months.<sup>1</sup> Compared with N1 disease, N2 disease is associated with worse survival after R0 resection (hazard ratio [HR], 1.65;  $p$  0.0001).<sup>2-4</sup> Additionally, increasing T stage (e.g., T1 to T4 R0) is associated with shorter 5-year OS and MST, from 50% and 60 months for T1N2 to 24% and 22 months for T4N2.<sup>5</sup> Interestingly, survival data differed substantially according to the region of the world analyzed, with the highest survival in North America (5-year OS, 42%; MST, 38.6 months) and the lowest in Europe (5-year OS, 22%; MST, 21.1 months).<sup>1,5</sup>

The revision of N descriptors in the 8<sup>th</sup> edition also uniquely includes differences in survival among various subsets of pN2 disease.<sup>3,6,7</sup> In fact, the current proposal separates single-station from multistation pN2 involvement, renaming these subcategories pN2a and pN2b, respectively.<sup>1,8</sup> The 5-year OS and MST for pN2a disease (49% and 57 months) were not statistically significantly different from those for pN1b disease (i.e., with multistation pN1 involvement; 50% and 60.9 months; HR, 1.04;  $p=0.67$ ). Conversely, outcomes for pN2b (38% and 40 months; HR, 1.47;  $p$  0.0001) were significantly different from those for pN2a. An important corollary to this analysis was the assessment of the “skip” phenomenon within the pN2a subset.<sup>1,9</sup> The absence of concomitant pN1 involvement in patients with pN2a disease (pN2a1; 54% and 70.9 months) was confirmed to be a favorable prognosticator,

compared with pN2a2 disease (i.e., with N1 involvement; 43% and 46 months; HR, 1.35;  $p$  0.0007).<sup>1</sup>

Irrespective of the known heterogeneity in survival among patients with pN2 disease across regions, a significant divergence was reported in the management of N2 disease between North America and Europe. In North America, induction therapy followed by surgery is favored, whereas in Europe, primary surgery followed by adjuvant treatment is the preferred choice.<sup>10,11</sup> An analysis of a combined Society of Thoracic Surgeons (STS) and European Society of Thoracic Surgeons (ESTS) database that included >78,000 surgical patients found a 2-fold higher use of lobectomy and pneumonectomy for N2 disease in the ESTS database, compared with the STS database.<sup>11</sup> This supports the argument that North American surgeons favor more-aggressive and invasive preoperative assessment of nodes (e.g., endobronchial ultrasound [EBUS], mediastinoscopy) along with induction therapy for biopsy-proven lymph nodes, whereas European surgeons are more inclined to proceed straight to surgery.

Primary surgical resection followed by adjuvant treatment is the customary approach for patients with cN0 status found to have so-called occult pN2 disease.<sup>7</sup> In these patients, pN2 disease is found after resection, even in the setting of radiologic assessment of lymph nodes and invasive mediastinal staging.<sup>12</sup> For this particular subset of patients, 5-year OS has been reported to be as high as 67%, with MST of 48.5 months.<sup>7,19</sup>

In this study, we investigate the relationship between two treatment approaches and survival outcomes among patients with pN2a1 NSCLC. As occult pN2a1 disease is believed to have the best survival among N2 subsets, we postulate that survival among patients with occult pN2a1 disease who undergo surgery first followed by adjuvant treatment (SF) could be used as a benchmark to measure outcomes in patients treated with neoadjuvant therapy followed by surgery (NS) who are found to have pN2a disease.

## METHODS

### Patient Selection

Following approval from our institutional review board, we performed a retrospective review of a prospectively maintained database to identify all patients who underwent surgical resection for lung cancer at our institution between January 2000 and April 2018. Patients with pathologically diagnosed NSCLC in a single station, without N1 involvement (pN2a1), who had R0 resection were included. Patients underwent systematic mediastinal lymphadenectomy including curative-intent mediastinal node dissection of at least 2 mediastinal stations. Patients with extranodal metastases, recurrent disease, or carcinoid tumors noted on radiologic imaging or final pathologic analysis were excluded.

### Patient Groups

Patients were separated into two groups on the basis of whether they underwent surgery before or after other treatments. The SF group included patients with true occult disease initially treated with surgery first followed by adjuvant treatment (cN0-pN2a1), and the NS group included patients treated with neoadjuvant therapy before surgery (cN2-pN2a1). All

treatments were administered at the discretion of the individual treatment team. Neoadjuvant chemotherapy dosing was left to the discretion of the patient's primary oncologist and consisted of a multimodal platinum-based regimen.

### Variables

Patient demographic and tumor characteristics—including age at surgery; forced expiratory volume in 1 second; diffusing capacity of the lung for carbon monoxide; sex; smoking status; pulmonary, cardiac, endocrine, renal, and other comorbidities; computed tomography (CT) imaging of the tumor; positron emission tomography (PET) scan characteristics of the tumor; approach (video-assisted thoracic surgery, yes or no); postoperative radiation therapy (yes or no); type of surgery (wedge/segment, lobectomy/bilobectomy, pneumonectomy); pathologic diagnosis (adenocarcinoma, squamous cell carcinoma, other); and clinical and pathologic stage—were collected. Differences in patient characteristics between the two groups were compared using the Wilcoxon rank-sum test for continuous variables and Fisher's exact test for categorical variables.

### Clinical Staging and Radiologic Assessment

Clinical staging and radiologic assessment of tumors were performed for each patient. Information on CT and PET scans were obtained for patients with available scans. Maximum standardized uptake value (SUV<sub>max</sub>) was obtained for the primary lesion (n=205) and nodal metastases (n=112). CT and PET scans were used to assess the initial clinical staging of nodes and potential extranodal metastases.

Invasive mediastinal staging was performed on the basis of any clinical suspicion of nodal disease from PET and CT findings. The majority of patients in the NS group received invasive mediastinal staging with either mediastinoscopy or EBUS (64%). The reason for not receiving invasive mediastinal staging (n=45, 36%) included imaging showing significant nodal disease (n=18), having started chemotherapy before seeing a surgeon (n=8), tumor size (n=7), being enrolled in a clinical trial (n=5), separate primaries (n=2), node location not amenable to invasive staging (n=2), multiple lung nodules (n=2), and functional status (n=1). Conversely, the majority of SF patients did not receive any invasive mediastinal staging (83%). All patients were clinically staged according to the 8<sup>th</sup> edition of the TNM staging system.

### Survival Analysis

The primary outcome of interest was disease-free survival (DFS), which was defined from the date of surgery until the date of recurrence or death without recurrence. The secondary outcome of interest was OS, which was defined from the date of surgery until the date of death. Patients were otherwise censored on the date of the last follow-up.

### Matching-Weights Procedure

To address potential covariate imbalance or selection bias between the SF group, comprising patients with true occult disease initially treated with surgery first followed by adjuvant treatment (cN0-pN2a1), and the NS group, comprising patients treated with neoadjuvant therapy before surgery (cN2-pN2a1), we applied a matching-weights approach to the

analysis of the primary and secondary endpoints. The matching-weights procedure (which is analogous to 1:1 pairwise propensity-score matching<sup>13</sup>) was used to generate a pseudopopulation in which the two groups have similar distributions of characteristics (i.e., balanced). Unlike the 1:1 propensity-score matching procedure, which excludes all unmatched patients, the matching-weights approach does not exclude patients; instead, it applies weights to each patient such that “unmatched” patients are down-weighted in the analyses. A logistic regression model was used to derive the odds of being in the NS group. The resulting patient-level matching-weights were based on the smaller of the predicted probabilities of being in the NS group, divided by the predicted probability of being assigned to the group the patient was actually in. Each patient contributed a fraction toward the overall analysis on the basis of the magnitude of the matching-weights.

The variables included in the logistic model were selected *a priori* on the basis of relevant clinical factors associated with the likelihood of being in the NS group. They include 20 preoperative, operative, and postoperative variables: sex, age, body mass index, ever smoker, pulmonary comorbidity, cardiac comorbidity, endocrine comorbidity, renal comorbidity, other comorbidity, forced expiratory volume in 1 second, diffusing capacity of the lung for carbon monoxide, SUV primary, surgical type (wedge/segmentectomy vs lobectomy/bilobectomy vs pneumonectomy), video-assisted thoracic surgery, laterality, preoperative radiation, postoperative radiation, histologic subtype (adenocarcinoma, squamous, other), pT stage, and year of surgery. The absolute standardized mean difference (ASMD) for each variable quantifies the performance of the matching-weights approach between patients in the SF group or the NS group. An ASMD  $\leq 0.1$  implies adequate covariate balance of the variable between the two groups.<sup>14</sup> The ASMD for each variable, before and after application of the matching-weights, are included in Supplementary Figure 1. Before applying the matching-weights, ASMD values were  $>0.1$  for 14 of 20 variables, confirming lack of balance between the two groups in terms of patient characteristics. After applying the matching-weights, the ASMD values were  $\leq 0.1$  across all variables considered, indicating successful balance across all clinically relevant variables between the two groups (Supplementary Figure 1). The distribution of the propensity scores before and after application of the matching-weights is presented as a mirror histogram for visual assessment of the success of the matching-weights approach; the mirror histogram shows good overlap in propensity scores between SF and NS patients after application of the matching-weights (Supplementary Figure 2).

For the matching-weights analyses, each patient contributes a fraction reflecting the matching-weights; therefore, the effective sample size may not be an integer. Survival was estimated using the Kaplan-Meier approach and compared between groups using a weighted log-rank test. The associations between survival endpoints and intervention (NS or SF) were quantified using Cox proportional hazards models. Patient-level matching-weights were incorporated in all analyses. Only patients with all variables were included in matching-weights comparisons.

As a secondary analysis, the above analyses were repeated among a larger cohort of all patients with single-station nodal involvement on final pathologic assessment, regardless of clinical N status, to not miss any clinically overstaged or understaged patients (cN0-cN2,

pN2a1; Supplementary Table 1). Specifically, the matching-weights procedure and subsequent Kaplan-Meier analyses of DFS and OS were performed between all patients with pN2a1 disease (n=266) who underwent SF (n=141) versus patients with pN2a1 disease who underwent NS (n=125) (Supplementary Figure 3 and 4).

All statistical tests were 2-sided, and  $p < 0.05$  was considered to indicate statistical significance. Statistical analyses were conducted using Stata 13.1 (StataCorp, College Station, TX) and R 3.5.1 (R Development Core Team, Vienna, Austria), including the *survival*, *rms* (Regression Modeling Strategies), and *riskRegression* R packages. The matching-weights procedure was performed with the *survey* and *tableone* R packages, downloaded in November 2018.

### Sensitivity Analysis

We present, as a sensitivity analysis, the ASMD derived from conventional 1-to-1 propensity-score matching procedures and the resulting Kaplan-Meier curves using the matched cohort (Supplementary Table 2, Supplementary Figure 5) to demonstrate the superior balancing performance of the matching-weights procedure used in this study. Propensity scores were computed as the conditional probability of NS using a logistic regression model with the same factors as in the matching-weights procedure. Propensity score-matched pairs were identified without replacement using a 1-to-1 nearest neighbor greedy matching algorithm with caliper width equal to 0.2 of the standard deviation of the logit of the propensity score, as recommended by Austin.<sup>15</sup> Unlike the matching-weights procedure, the 1-to-1 propensity score-matching procedure was less successful, as 12 of 20 variables were still unbalanced between the two groups.

## RESULTS

### Demographic Characteristics and Comorbidities

A total of 227 patients with pN2a1 disease met the inclusion criteria for the primary analysis: 121 in the SF group and 106 in the NS group (Figure 1). The distribution between the two groups before the propensity-score matching-weights procedure is presented in Table 1. The proportion of males was higher in the NS group than in the SF group (48% vs 38%). Patients were also younger in the NS group than in the SF group (median, 65 [interquartile range {IQR}, 57-73] vs 68 [60-74] years). Patients in the NS group had slightly higher forced expiratory volume in 1 second (median, 88% [IQR, 77%-102%] vs 82% [73%-96%] in SF) (Table 1).

### Tumor Assessment

Before the propensity-score matching weights procedure, the SF group had smaller primary tumors on radiologic assessment (median, 2.5 cm [IQR, 1.5-3.5] vs 3.1 cm [2.2-5.0] in NS), with lower fluorodeoxyglucose uptake (median SUV, 6.4 [IQR, 3.3-11.9] vs 9.3 [4.8-14.4] in NS). Nodal uptake on PET imaging was higher in the NS group than in the SF group (median SUV, 7.0 [IQR, 4.5-10.0] vs 3.2 [2.7-5.3]). The intraoperatively assessed total number of nodes was lower in the SF group (median, 10 [IQR, 5-17]) than in the NS group

(12 [7-19]). The number of nodal stations assessed intraoperatively was also lower in the SF group (median, 4 [IQR, 3-5]; mean, 3.9) than in the NS group (4 [4-5]; mean, 4.5).

### Treatment Characteristics

Before the matching-weights procedure, patients in the SF group were more likely to undergo minimally invasive surgery (38% vs 13%) and less likely to undergo postoperative radiotherapy (34% vs 61%) than the NS group. Adjuvant platinum-based multimodal chemotherapy was given to 60% of patients in the SF group, compared with only 1% of patients in the NS group. Patients in the NS group were less likely to receive wedge resection/segmentectomy (16% vs 21%) and more likely to receive lobectomy/bilobectomy (81% vs 77%), compared with patients in the SF group.

### Disease-Free Survival

DFS was not significantly different between groups both before and after application of the matching-weights procedure (Figure 2). Before the matching-weights procedure, 5- and 10-year DFS (95% confidence interval [CI]) were 41% (33%-52%) and 20% (13%-32%) for the SF group versus 32% (23%-43%) and 21% (13%-34%) for the NS group (log-rank  $p=0.231$ ; NS vs reference of SF: hazard ratio [HR], 1.22; 95% CI, 0.88-1.68). After the matching-weights procedure, 5- and 10-year DFS (95% CI) were 45% (31%-66%) and 27% (13%-55%) for the SF group versus 26% (14%-48%) and 21% (9%-47%) for the NS group (log-rank  $p=0.056$ ; NS vs reference of SF: HR, 1.61; 95% CI, 0.98-2.65). Thirty of the 121 patients in the SF group had a distant recurrence, while 10 had a local recurrence (mediastinal lymph nodes and ipsilateral lung/thorax). Forty-one of the 106 patients in the NS group had a distal recurrence, while 10 had a local recurrence (Supplementary Table 3).

### Overall Survival

Similarly, OS was not significantly different between groups both before and after the matching-weights procedure (Figure 3). Before the matching-weights procedure, 5-year and 10-year OS (95% CI) were 52% (43%-63%) and 22% (14%-36%) for the SF group versus 49% (40%-62%) and 23% (14%-40%) for the NS group (log-rank  $p=0.827$ ; NS vs reference of SF: HR, 1.04; 95% CI, 0.73-1.47). After the matching-weights procedure, 5- and 10-year OS (95% CI) were 49% (34%-71%) and 30% (15%-59%) for the SF group versus 43% (28%-65%) and 20% (7%-58%) for the NS group (log-rank  $p=0.428$ ; NS vs reference of SF: HR, 1.24; 95% CI, 0.67-2.28).

### All Patients with pN2a1: SF versus NS

A secondary analysis of the expanded cohort among all patients with pN2a1 disease ( $n=266$ ) was performed between SF ( $n=141$ ) and NS ( $n=125$ ). There were no statistically significant differences in DFS or OS between the groups both before and after the matching-weights procedure (Figure 4). After the matching-weights procedure, 5- and 10-year DFS (95% CI) were 44% (31%-63%) and 29% (16%-54%) for the SF group versus 30% (18%-48%) and 18% (8%-41%) for the cN2 NS group (log-rank  $p=0.056$ ; NS vs reference of SF: HR, 1.52; 95% CI, 0.99-2.34), and 5- and 10-year OS (95% CI) were 49% (36%-68%) and 33%

(19%-58%) for the SF group versus 46% (33%-65%) and 22% (10%-49%) for the cN2 NS group (log-rank  $p=0.328$ ; NS vs reference of SF: HR, 1.26; 95% CI, 0.79-1.99).

## DISCUSSION

In the absence of conclusive evidence to support a unique and definitive treatment strategy for N2 disease, a heterogeneity of outcomes must be accepted.<sup>5</sup> At one extreme, recently published results from the Pacific Trial—which demonstrated the efficacy of immunotherapy after chemoradiotherapy for unresectable locally advanced lung cancer—have been interpreted as a requiem for surgery for stage IIIA NSCLC.<sup>16,17</sup> Nevertheless, in previous randomized trials of patients with locally advanced lung cancer, OS was clearly better among patients managed with multimodality treatment that included surgery than among patients treated with chemotherapy alone.<sup>18</sup> In addition, in the reporting on the Pacific Trial, the researchers do not acknowledge the prognostic distinction between subsets of N2 disease—in particular, the difference between single-station and multistation N2 disease is not noted.<sup>17</sup>

At the other extreme, survival following SF for unsuspected occult pN2 disease may overlap survival for pN2a and even pN1 disease, and survival for occult pN2 disease is undisputedly the best for the entire N2 group.<sup>7,19,20</sup> One caveat when considering occult pN2 disease is that we may be focusing on a subset of patients with limited disease and, possibly, a peculiar biology. In fact, a previous study from our institution has demonstrated that, in 16% of patients, occult N2 disease follows lymphatic routes beyond the usual drainage pathways.<sup>21</sup>

The interrogation of large databases to investigate NS versus SF for cN2 disease has not answered the question of which strategy is superior.<sup>22–24</sup> In fact, results from the Surveillance, Epidemiology, and End Results and STS databases showed substantial equipoise between approaches. However, when the National Cancer Database was queried, survival outcomes were inconclusive, as one study showed survival was better for patients treated with NS,<sup>25</sup> while another showed no statistically significant differences.<sup>23</sup> The slight advantage associated with NS over SF was confirmed in a recently published network study in which the HR for NS was 1.14.<sup>26</sup>

The aim of our study was to investigate the relationship between NS and survival among patients with occult pN2a1 disease and to compare that with the relationship between SF and survival, as SF has been shown to be associated with the most-favorable outcomes.<sup>27</sup> Although we observed some evidence of worse DFS among those treated with NS, compared with SF, differences between groups did not meet the conventional level of statistical significance. We have demonstrated through this retrospective analysis that, regardless of preoperative treatment, long-term survival is similar when pN2a1 disease is found postoperatively.

Our study has several limitations. The study features a retrospective design and was performed at a single institution, which may have resulted in selection bias and limited generalizability. All patients were treated at the discretion of the practicing surgeon and treatment team. As this is a retrospective study spanning 18 years, not all patients underwent



the same preoperative workup with PET scans or invasive mediastinal staging. Also, we were unable to fully gather the number of patients with clinical N2 disease who received chemotherapy but never underwent a surgical intervention, as well as patients who were downstaged to N0 after resection. Next, while we are diligent regarding intraoperative lymph node assessment, labeling, and handling of lymph nodes, per institutional standards, there is always the chance of mislabeling or miscounting lymph nodes. This has the potential to make unrecognized N1 disease a confounder in this retrospective study. Last, systemic therapy was not constant over time, and patients received different systemic therapy depending on the practicing team.

Important factors for producing a reliable analysis include the accuracy of staging and the completeness of the database.<sup>23</sup> In this setting, we consider it crucial that mediastinal staging, such as EBUS, performed by a thoracic surgeon is used liberally. In addition, at our institution, we conduct a weekly review of the individual stage for each patient during an *ad hoc* staging conference. To ensure accuracy, the data from all surgical patients are collectively reviewed by the faculty and amended real-time in the database.

The literature includes countless studies on N2 disease based on either the clinical or the pathologic N2 subsets (cN2 and pN2).<sup>23</sup> These studies, as well as the current one, implicitly carry the risk of including downstaged patients (i.e., cN2 patients rendered pN0 after treatment) or missing upstaged patients (i.e., cN0 patients who become pN2). In this setting, we performed an analysis of NS versus SF for patients with cN2-pN2a1 compared with cN0-pN2a1 disease and observed no statistically significant difference in OS and DFS.

## CONCLUSIONS

While our results lend support to the concept that adequate NS regimens can yield DFS and OS comparable to the benchmark survival achieved with SF in patients with occult pN2a1 disease, they also demonstrate that SF can be an acceptable treatment regimen for patients with pN2a1 disease. In addition, we have shown that, regardless of the timing of chemotherapy (NS or SF), patients found to have pN2a1 disease have similar long-term outcomes. Regardless, to truly identify the ideal timing of chemotherapy for patients with lung cancer with preresectional histologically confirmed N2a disease, a randomized controlled trial would be necessary. Figure 5 shows a possible outline of future studies in this area. In the meantime, the SF strategy offers reduced time from diagnosis to definitive surgical intervention and the ability to perform a more rapid and feasible surgical procedure by minimally invasive techniques, as the surgeon is less likely to encounter difficult adhesions and altered anatomical planes, especially in the hilum and the mediastinum, which are usually found after induction therapy. Conversely, the NS strategy offers the advantages of reduced micrometastatic disease and an increased likelihood of patients completing systemic treatment. This is especially true for larger tumors with higher metabolic activity, as determined by PET scan. In addition, patients with pN2a disease may be further downstaged, with an attendant benefit in survival.

In conclusion, given the lack of significant survival differences between SF and NS (Figure 6), pN2a seems to be the perfect subset in which to verify the best timing for chemotherapy

in the context of a randomized trial, preferably in a multicenter setting, given the anticipated low number of patients in this subset at a single institution. In this setting, an estimate of the trial power could be determined by taking into account (a) the difference in DFS between NS for presectional known pN2a1 disease and SF for occult pN2a1 disease in the present study and (b) the difference in DFS between SF for presectional known pN2a1 disease and SF for occult pN2a1 disease. The latter data could be generated at international centers, where SF is considered the standard of care for single-station N2 disease.<sup>11</sup>

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

**Funding:** This work was supported, in part, by NIH/NCI Cancer Center Support Grant P30 CA008748.

**COI statement:** The authors have the disclosures listed below.

G.R. is a consultant for Scanlan International.

J.M.I. is a consultant for LumaCyte.

M.J.B. is a consultant for AstraZeneca Pharmaceuticals.

V.W.R. receives grant support from Stand Up To Cancer, Memorial Sloan Kettering Cancer Center, and Genelix.

D.R.J. serves as a senior medical advisor for Diffusion Pharmaceuticals and a consultant for Merck and AstraZeneca.

## Glossary of Abbreviations

<b>ASMD</b>	absolute standardized mean difference
<b>CI</b>	confidence interval
<b>CT</b>	computed tomography
<b>DFS</b>	disease-free survival
<b>EBUS</b>	endobronchial ultrasound
<b>ESTS</b>	European Society of Thoracic Surgeons
<b>HR</b>	hazard ratio
<b>MST</b>	median survival time
<b>N</b>	nodal
<b>N2a</b>	single-station N2
<b>NS</b>	neoadjuvant therapy followed by surgery
<b>NSCLC</b>	non-small cell lung cancer

<b>OS</b>	overall survival
<b>PET</b>	positron emission tomography
<b>pN2</b>	pathologic N2
<b>pN2a</b>	pathologic single-station N2
<b>pN2a1</b>	pathologic single-station N2 (without any N1 involvement)
<b>SF</b>	surgery first followed by adjuvant treatment
<b>STS</b>	Society of Thoracic Surgeons
<b>SUVmax</b>	maximum standardized uptake value
<b>TNM</b>	tumor-node-metastasis

## REFERENCES

1. Asamura H, Chansky K, Crowley J, Goldstraw P, Rusch VW, Vansteenkiste JF, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: proposals for the revision of the N descriptors in the forthcoming 8th edition of the TNM classification for lung cancer. *J Thorac Oncol.* 2015;10(12):1675–84. [PubMed: 26709477]
2. Farjah F, Lou F, Sima C, Rusch VW, Rizk NP. A prediction model for pathologic N2 disease in lung cancer patients with a negative mediastinum by positron emission tomography. *J Thorac Oncol.* 2013;8(9):1170–80. [PubMed: 23945387]
3. Hanagiri T, Takenaka M, Oka S, Shingematsu Y, Nagata Y, Shimokawa H, et al. Clinical significance in the number of involved lymph nodes in patients that underwent surgery for pathological stage III-N2 non-small cell lung cancer. *J Cardiothorac Surg.* 2011;6:144. [PubMed: 22027105]
4. Farjah F, Backhus LM, Varghese TK, Manning JP, Cheng AM, Mulligan MS, Wood DE. External validation of a prediction model for pathologic N2 among patients with a negative mediastinum by positron emission tomography. *J Thorac Dis.* 2015;7(4):576–84. [PubMed: 25973222]
5. Rami-Porta R, Ball D, Crowley J, Giroux DJ, Jett J, Travis WD, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the T descriptors in the forthcoming (seventh) edition of the TNM classification for lung cancer. *J Thorac Oncol.* 2007;2(7):593–602. [PubMed: 17607114]
6. Hishida T, Yoshida J, Ohe Y, Aokage K, Ishii G, Nagai K. Surgical outcomes after initial surgery for clinical single-station N2 non-small-cell lung cancer. *Jpn J Clin Oncol.* 2014;44(1):85–92. [PubMed: 24203815]
7. Cho HJ, Kim SR, Kim HR, Han JO, Kim YH, Kim DK, et al. Modern outcome and risk analysis of surgically resected occult N2 non-small cell lung cancer. *Ann Thorac Surg.* 2014;97(6):1920–5. [PubMed: 24768044]
8. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, 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. [PubMed: 26762738]
9. Nicholson AG, Tsao MS, Travis WD, Patil DT, Galateau-Salle F, Marino M, et al. Eighth edition staging of thoracic malignancies: implications for the reporting pathologist. *Arch Pathol Lab Med.* 2018;142(5):645–61. [PubMed: 29480761]
10. Veeramachaneni NK, Feins RH, Stephenson BJ, Edwards LJ, Fernandez FG. Management of stage IIIA non-small cell lung cancer by thoracic surgeons in North America. *Ann Thorac Surg.* 2012;94(3):922–6; discussion 926–8. [PubMed: 22742842]

11. Rocco G, Nason K, Brunelli A, Varela G, Waddell T, Jones DR. Management of stage IIIA (N2) non-small cell lung cancer: a transatlantic perspective. *J Thorac Cardiovasc Surg.* 2016;151(5):1235–8. [PubMed: 26997100]
12. Szlubowski A, Zielinski M, Soja J, Filarecka A, Orzechowski S, Pankowski J, et al. Accurate and safe mediastinal restaging by combined endobronchial and endoscopic ultrasound-guided needle aspiration performed by single ultrasound bronchoscope. *Eur J Cardiothorac Surg.* 2014;46(2):262–6. [PubMed: 24420366]
13. Li L, Greene T. A weighting analogue to pair matching in propensity score analysis. *Int J Biostat.* 2013;9(2):215–34. [PubMed: 23902694]
14. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Res.* 2011;46(3):399–424.
15. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm Stat.* 2011;10(2):150–61. [PubMed: 20925139]
16. Antonia SJ, Villegas A, Daniel D, Vicente D, Murakami S, Hui R, et al. Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. *N Engl J Med.* 2017;377(20):1919–29. [PubMed: 28885881]
17. Pass HI. PACIFIC: Time for a surgical IIIA uprising. *J Thorac Cardiovasc Surg.* 2018;156(3):1249–54. [PubMed: 30017445]
18. McElnay PJ, Choong A, Jordan E, Song F, Lim E. Outcome of surgery versus radiotherapy after induction treatment in patients with N2 disease: systematic review and meta-analysis of randomised trials. *Thorax.* 2015;70(8):764–8. [PubMed: 25967753]
19. Honguero Martinez AF, Garcia Jimenez MD, Garcia Vicente A, Genoves Crespo M, Rodriiguez Ortega CR, Lazaro Sahuguillo M, et al. Is the prognosis of occult N2 disease similar to that of positive positron emission tomography-computed tomography (PET/CT) scan single-station N2 disease in patients with non-small cell lung cancer treated by surgical resection? *Rev Espanola Med Nucl E Imagen Mol.* 2017;36(6):350–5.
20. Kim MS, Lee H-S, Lee JM, Zo JI, Lee GK, Nam B-H. Prognostic value of single nodal zone metastasis in non-small-cell lung cancer. *Eur J Cardiothorac Surg.* 2010;38(4):491–7. [PubMed: 20399672]
21. Bille A, Woo KM, Ahmad U, Rizk NP, Jones DR. Incidence of occult pN2 disease following resection and mediastinal lymph node dissection in clinical stage I lung cancer patients. *Eur J Cardiothorac Surg.* 2017;51(4):674–9. [PubMed: 28200091]
22. Boffa D, Fernandez FG, Kim S, Kosinski A, Onaitis MW, Cowper P, et al. Surgically managed clinical stage IIIA-clinical N2 lung cancer in the Society of Thoracic Surgeons Database. *Ann Thorac Surg.* 2017;104(2):395–403. [PubMed: 28527969]
23. Boffa DJ, Hancock JG, Yao X, Goldberg S, Rosen JE, Kim AW, et al. Now or later: evaluating the importance of chemotherapy timing in resectable stage III (N2) lung cancer in the National Cancer Database. *Ann Thorac Surg.* 2015;99(1):200–8. [PubMed: 25440272]
24. Yendamuri S, Dhillon SS, Groman A, Dy G, Dexter E, Picone A, et al. Effect of the number of lymph nodes examined on the survival of patients with stage I non-small cell lung cancer who undergo sublobar resection. *J Thorac Cardiovasc Surg.* 2018;156(1):394–402. [PubMed: 29709364]
25. Koshy M, Fedewa SA, Malik R, Ferguson MK, Vigeswaran WT, Feldman L, et al. Improved survival associated with neoadjuvant chemoradiation in patients with clinical stage IIIA(N2) non-small-cell lung cancer. *J Thorac Oncol.* 2013;8(7):915–22. [PubMed: 23608815]
26. Zhao Y, Wang W, Liang H, Yang CJ, D'Amico T, Ng CSH, et al. The optimal treatment for stage IIIA-N2 non-small-cell lung cancer: a network meta-analysis. *Ann Thorac Surg.* 2019;107(6):1866–75. [PubMed: 30557543]
27. Romesser PB, Bardash Y, Buonocore D, Chافت JE, Huang J, Jones DR, et al. Outcomes of stage III NSCLC with occult primary vs. known primary lesions. *Lung Cancer.* 2019;127:34–36. [PubMed: 30642548]

**Central Message:**

No statistically significant difference has been observed between SF for occult single-station N2 (pN2a1) NSCLC and NS for known pathologic single-station N2.

Author Manuscript

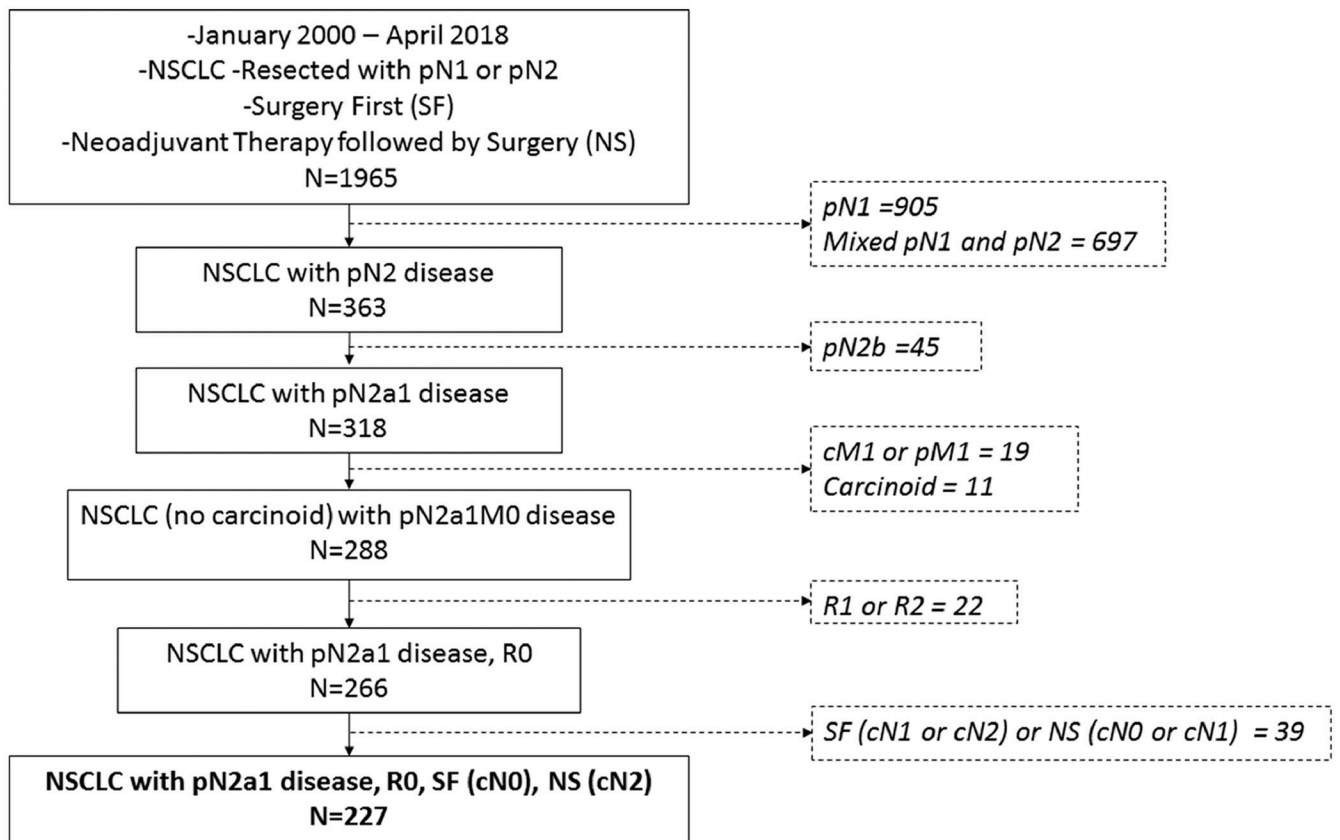
Author Manuscript

Author Manuscript

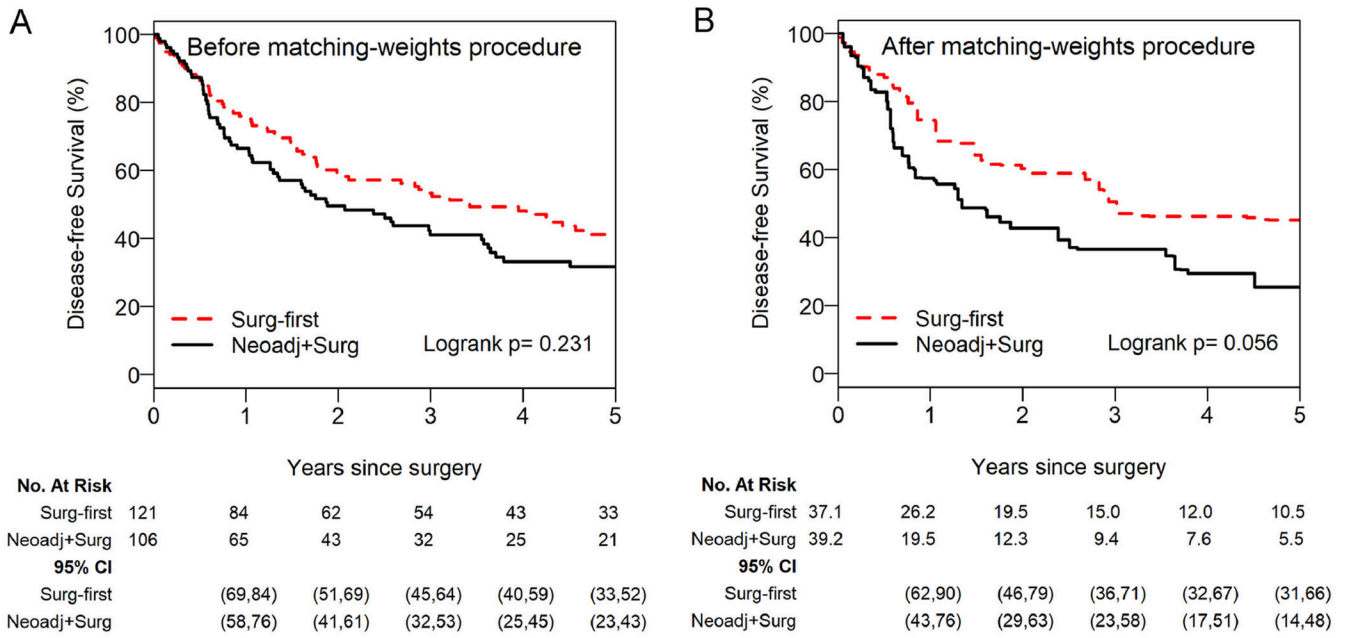
Author Manuscript

**Perspective Statement:**

Surgery first for occult single-station N2 (pN2a1) portends the best prognosis among N2 subsets. In our series, no significant difference in DFS was noted between surgery first—including occult pN2a1—and neoadjuvant therapy followed by surgery. Building on our findings, further investigation of surgery first for known N2a1 versus occult pN2a1 disease could power a clinical trial focused on N2a NSCLC.

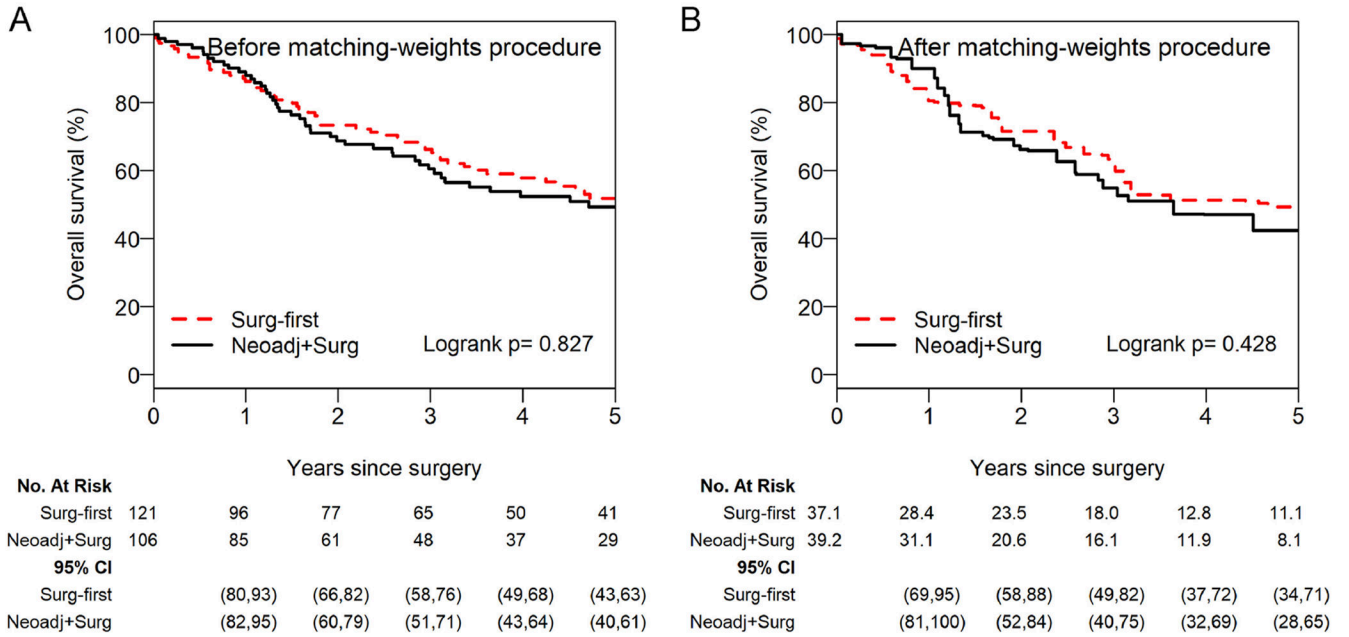


**Figure 1.** CONSORT diagram for study inclusion. NSCLC, non-small cell lung cancer.

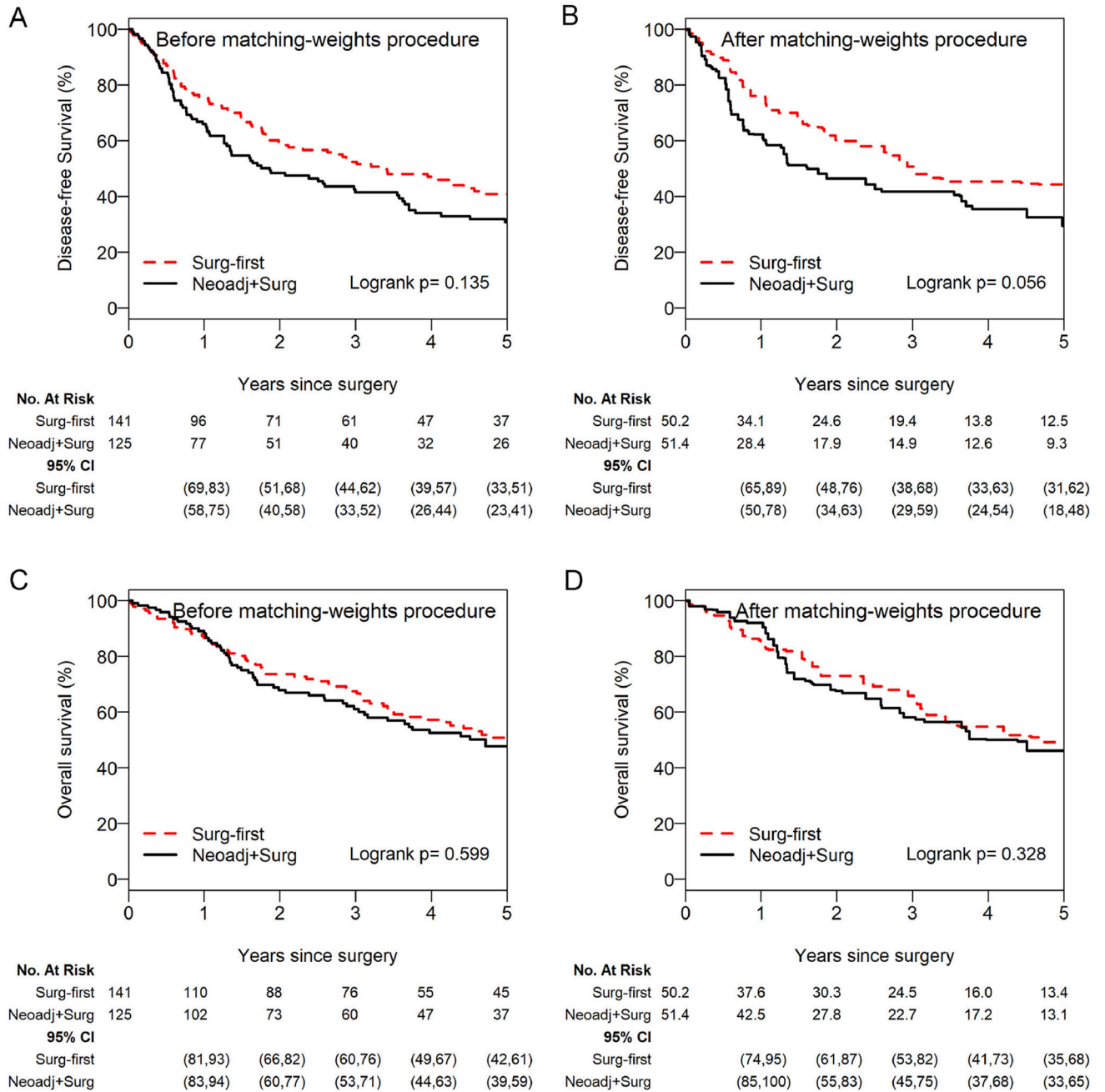


**Figure 2.** Disease-free survival showed no significant differences both before (A) and after (B) applying the matching-weights procedure for patients with cN0-pN2a1 non-small cell lung cancer who underwent NS versus patients who underwent SF for cN2-pN2a1 disease. Cox model after the matching-weights procedure: (NS vs reference of SF) HR, 1.61 (95% CI, 0.98-2.65). For the matching-weights analyses, each patient contributes a fraction reflecting the matching-weights; therefore, the effective sample size (numbers at risk) may not be an integer. The number of patients at risk and 95% CI estimates corresponding to each time point are presented below the figure for each group. CI, confidence interval; HR, hazard ratio; NS, neoadjuvant therapy followed by surgery; SF, surgery first.





**Figure 3.** Overall survival was not significantly different between patients with cN0-pN2a1 non-small cell lung cancer who underwent NS and patients who underwent SF for cN2-pN2a1 before (A) and after (B) applying the matching-weights procedure. Cox model after the matching-weights procedure: (NS vs reference of SF) HR, 1.24 (95% CI, 0.67-2.28). For the matching-weights analyses, each patient contributes a fraction reflecting the matching-weights; therefore, the effective sample size (numbers at risk) may not be an integer. The number of patients at risk and 95% CI estimates corresponding to each time point are presented below the figure for each group. CI, confidence interval; HR, hazard ratio; NS, neoadjuvant therapy followed by surgery; SF, surgery first.



**Figure 4.** Disease-free survival (A, B) and overall survival (C, D) were not significantly different between patients who underwent NS and patients who underwent SF among all patients with pathologic N2a1 non-small cell lung cancer before and after the matching-weights procedure. Disease-free survival: Cox model after the matching-weights procedure: (NS vs reference of SF) HR, 1.52 (95% CI, 0.99-2.34). Overall survival: Cox model after the matching-weights procedure: (NS vs reference of SF) HR, 1.26 (95% CI, 0.79-1.99). For the matching-weights analyses, each patient contributes a fraction reflecting the matching-

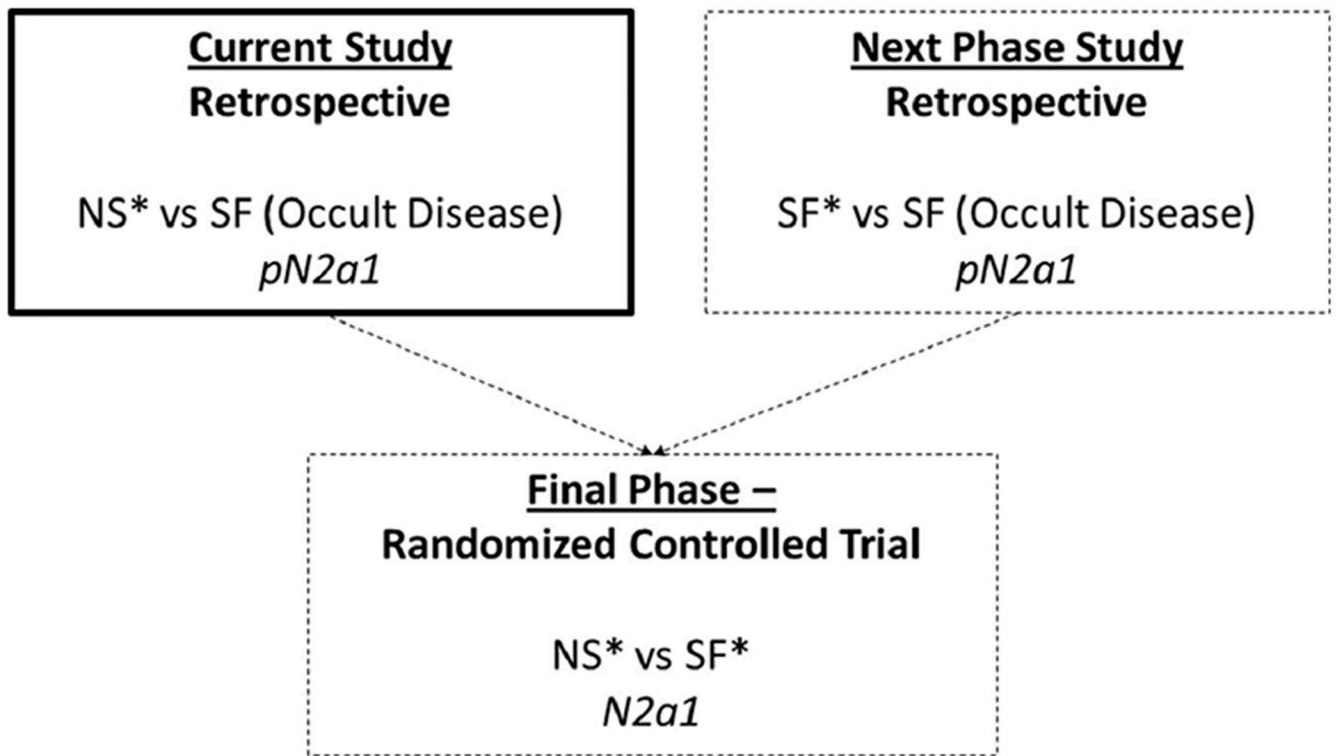
weights; therefore, the effective sample size (numbers at risk) may not be an integer. The number of patients at risk and 95% CI estimates corresponding to each time point are presented below the figure for each group. CI, confidence interval; HR, hazard ratio; NS, neoadjuvant therapy followed by surgery; SF, surgery first.

Author Manuscript

Author Manuscript

Author Manuscript

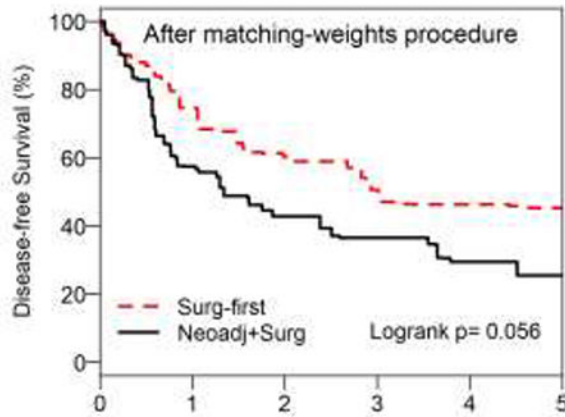
Author Manuscript



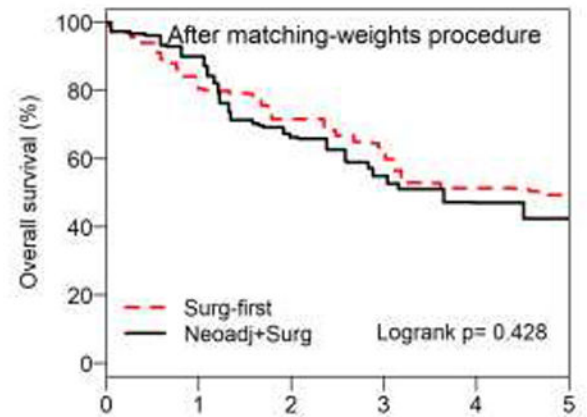
**Figure 5.**

Proposal for studies to help answer whether surgery first (SF) or neoadjuvant therapy followed by surgery (NS) results in better outcomes for single-station N2 disease.

\*Histologically confirmed N2 disease.



No. At Risk	Years since surgery					
	0	1	2	3	4	5
Surg-first	37.1	26.2	19.5	15.0	12.0	10.5
Neoadj+Surg	39.2	19.5	12.3	9.4	7.6	5.5
<b>95% CI</b>						
Surg-first		(62,90)	(46,79)	(36,71)	(32,67)	(31,66)
Neoadj+Surg		(43,76)	(29,63)	(23,58)	(17,51)	(14,48)



No. At Risk	Years since surgery					
	0	1	2	3	4	5
Surg-first	37.1	28.4	23.5	18.0	12.8	11.1
Neoadj+Surg	39.2	31.1	20.6	16.1	11.9	8.1
<b>95% CI</b>						
Surg-first		(69,95)	(58,88)	(49,82)	(37,72)	(34,71)
Neoadj+Surg		(81,100)	(52,84)	(40,75)	(32,69)	(28,65)

**Central Picture:**

DFS and OS after the matching-weights procedure among patients with pN2a1 NSCLC.

**Table 1.**

Demographic and tumor characteristics before and after the matching-weights procedure

Variable	Before matching-weights procedure			After matching-weights procedure		
	Surgery First (n=121)	Neoadjuvant Therapy + Surgery (n=106)	AS MD	Surgery First (n=37.1)	Neoadjuvant Therapy + Surgery (n=39.2)	AS MD
Male	46 (38.0)	51 (48.1)	0.205	17.3 (46.6)	17.3 (44.2)	0.047
Age, years	68.0 (60.0-74.0)	65.0 (57.3-72.8)	0.369	68.0 (64.0-73.4)	70.0 (61.0-74.0)	0.047
BMI	26.3 (23.5-29.9)	28.2 (24.7-32.0)	0.214	27.2 (24.9-30.3)	27.8.0 (24.6-30.5)	0.050
Smoking, yes	98 (81.0)	95 (89.6)	0.246	32.9 (88.5)	33.9 (86.6)	0.060
Pulmonary comorbidity, yes	42 (34.7)	37 (34.9)	0.004	14.6 (39.4)	15.6 (39.8)	0.007
Cardiac comorbidity, yes	71 (58.7)	59 (55.7)	0.061	22.5 (60.7)	25.0 (63.8)	0.065
Endocrine comorbidity, yes	17 (14.0)	11 (10.4)	0.112	2.9 (7.8)	3.5 (9.0)	0.045
Renal comorbidity, yes	4 (3.3)	3 (2.8)	0.028	0.9 (2.4)	1.1 (2.8)	0.024
Other comorbidity, yes	7 (5.8)	7 (6.6)	0.034	3.1 (8.5)	2.9 (7.5)	0.038
FEV1, %	82.0 (72.5-96.0)	88.0 (77.0-102.0)	0.234	86.0 (77.2-102.2)	86.0 (75.0-104.2)	0.005
DLCO, %	72.0 (58.0-83.0)	72.0 (61.0-84.0)	0.096	70.0 (56.2-78.0)	70.0 (56.0-78.8)	0.033
SUV primary	6.4 (3.3-11.9)	9.3 (4.8-14.4)	0.322	7.2 (3.3-12.5)	8.1 (4.3-13.5)	0.025
Surgery type			0.120			0.033
Wedge/segmentectomy	25 (20.7)	17 (16.0)		6.2 (16.8)	6.5 (16.5)	
Lobectomy/bilobectomy	93 (76.9)	86 (81.1)		29.7 (80.1)	31.7 (80.9)	
Pneumonectomy	3 (2.5)	3 (2.8)		1.2 (3.1)	1.0 (2.6)	
Video-assisted thoracic surgery, yes	46 (38.0)	14 (13.2)	0.593	8.5 (23.0)	8.9 (22.7)	0.007
Laterality, left	47 (38.8)	19 (17.9)	0.477	7.9 (21.4)	8.4 (21.5)	0.002
Preoperative radiotherapy, yes	0 (0)	7 (6.6)	0.376	0 (0)	0 (0)	<0.001
Pathologic diagnosis			0.397			0.033
Adenocarcinoma	102 (84.3)	72 (67.9)		29.1 (78.3)	30.3 (77.5)	
Squamous	12 (9.9)	19 (17.9)		6.2 (16.7)	6.6 (16.8)	
Other	7 (5.8)	15 (14.2)		1.9 (5.0)	2.2 (5.7)	
Pathologic T stage			0.592			0.078
0	0 (0)	8 (7.5)		0 (0)	0 (0)	
1	59 (48.8)	53 (50.0)		18.5 (49.9)	19.7 (50.3)	
2	41 (33.9)	24 (22.6)		11.3 (30.4)	12.8 (32.7)	
3	10 (8.3)	18 (17.0)		4.9 (13.3)	4.7 (11.9)	
4	11 (9.1)	3 (2.8)		2.4 (6.4)	2.0 (5.1)	
Postoperative radiotherapy, yes	41 (33.9)	65 (61.3)	0.571	22.2 (40.1)	24.0 (38.9)	0.023
Year of surgery			0.076			0.066
2003-2004	31 (25.6)	28 (26.4)		7.5 (20.2)	7.1 (18.2)	
2005-2009	30 (24.8)	24 (22.6)		8.1 (21.8)	8.9 (22.6)	
2010-2014	39 (32.2)	33 (31.1)		13.6 (36.7)	14.0 (35.9)	

Variable	Before matching-weights procedure			After matching-weights procedure		
	Surgery First (n=121)	Neoadjuvant Therapy + Surgery (n=106)	AS MD	Surgery First (n=37.1)	Neoadjuvant Therapy + Surgery (n=39.2)	AS MD
<b>2015 or after</b>	21 (17.4)	21 (19.8)		7.9 (21.3)	9.1 (23.3)	

Data are no. (%) or median (interquartile range). ASMD, absolute standardized mean difference; BMI, body mass index; CT, computed tomography; DLCO, diffusing capacity of the lung for carbon monoxide; FEV1, forced expiratory volume in 1 second; SUV, standardized uptake value. ASMD < 0.1 indicates adequate balance between the two groups. After the matching-weights procedure, each patient contributes a fraction reflecting the matching-weights; therefore, the effective sample size may not be an integer.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript