



HAL
open science

Sleeve lobectomy may provide better outcomes than pneumonectomy for non-small cell lung cancer. A decade in a nationwide study

Pierre-Benoit Pagès, Pierre Mordant, Stéphane Renaud, Laurent Brouchet, Pascal-Alexandre Thomas, Marcel Dahan, Alain Bernard

► To cite this version:

Pierre-Benoit Pagès, Pierre Mordant, Stéphane Renaud, Laurent Brouchet, Pascal-Alexandre Thomas, et al.. Sleeve lobectomy may provide better outcomes than pneumonectomy for non-small cell lung cancer. A decade in a nationwide study. *Journal of Thoracic and Cardiovascular Surgery*, 2017, 153 (1), pp.184+. 10.1016/j.jtcvs.2016.09.060 . hal-01489668

HAL Id: hal-01489668

<https://hal.science/hal-01489668v1>

Submitted on 26 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Sleeve lobectomy may provide better outcomes than pneumonectomy for non-small cell lung cancer. A decade in a nationwide study

Pierre-Benoit Pagès, MD, PhD, Pierre Mordant, MD, PhD, Stéphane Renaud, MD, MSc, Laurent Brouchet, MD, PhD, Pascal-Alexandre Thomas, MD, PhD, Marcel Dahan, MD, PhD, Alain Bernard, MD, PhD

PII: S0022-5223(16)31275-2

DOI: [10.1016/j.jtcvs.2016.09.060](https://doi.org/10.1016/j.jtcvs.2016.09.060)

Reference: YMTC 10927

To appear in: *The Journal of Thoracic and Cardiovascular Surgery*

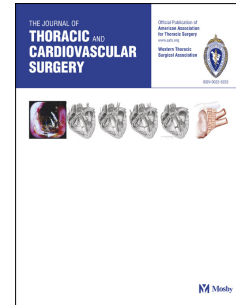
Received Date: 26 January 2016

Revised Date: 15 September 2016

Accepted Date: 23 September 2016

Please cite this article as: Pagès P-B, Mordant P, Renaud S, Brouchet L, Thomas P-A, Dahan M, Bernard A, for the EPITHOR project (French Society of Thoracic and Cardiovascular Surgery), Sleeve lobectomy may provide better outcomes than pneumonectomy for non-small cell lung cancer. A decade in a nationwide study, *The Journal of Thoracic and Cardiovascular Surgery* (2016), doi: 10.1016/j.jtcvs.2016.09.060.

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.



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Sleeve lobectomy may provide better outcomes than pneumonectomy for non-small cell lung cancer. A decade in a nationwide study

Pierre-Benoit Pagès, MD, PhD¹, Pierre Mordant, MD, PhD², Stéphane Renaud, MD, MSc³, Laurent Brouchet, MD, PhD⁴, Pascal-Alexandre Thomas, MD, PhD⁵, Marcel Dahan, MD, PhD⁶ and Alain Bernard, MD, PhD¹ for the EPITHOR project (French Society of Thoracic and Cardiovascular Surgery).

1- CHU Dijon, Hôpital du Bocage, 14 rue Gaffarel, 21079 Dijon

2- AP-HP, Hôpital Bichat-Claude Bernard, 46 Rue Henri Huchard, 75018 Paris

3- CHU Strasbourg, Hôpital Civil, 1 Place de l'hôpital, 67200 Strasbourg

4- CHU Toulouse, Hôpital Larrey, 24 chemin de Pouvoirville, 31000 Toulouse

5- CHU Marseille, Hôpital Nord, Chemin des Bourrely, 13915 Marseille cedex

All in France

Word count: 3688 Words

Abstract: 244 words

Corresponding author:

Pierre-Benoit Pagès, MD, PhD

Department of Thoracic and Cardiovascular Surgery

CHU Dijon, Hôpital du Bocage, 14 rue Gaffarel, BP 77908, 21079 Dijon, France

Tel.: +33 380293352

Fax: +33 380293385

E-mail: pierrebenoit.pages@chu-dijon.fr

Conflicts of interest and source of funding

29 The other authors have no disclosure to declare for this work.

ACCEPTED MANUSCRIPT

30 **Abstract**

31 **Introduction:** Whenever feasible, sleeve lobectomy is recommended to avoid pneumonectomy for
32 lung cancer, but these guidelines are based on limited retrospective series. The aim of our study was to
33 compare outcomes following sleeve lobectomy and pneumonectomy using data from a national
34 database.

35
36 **Methods:** From 2005 to 2014, 941 sleeve lobectomy and 5318 pneumonectomy patients were
37 recorded in the French database Epithor. Propensity score was generated with 15 pretreatment
38 variables and used to create balanced groups with matching (794 matches) and inverse probability of
39 treatment weighting (standardized difference was 0 for matching, and 0.0025 after weighting). Odds
40 ratio of postoperative complications and mortality, hazard ratio for overall survival and disease-free
41 survival were calculated using propensity adjustment techniques and a sensitivity analysis.

42
43 **Results:** Postoperative mortality after sleeve resection was similar to that after pneumonectomy
44 (Matching OR 1.24, $p=0.4$; weighting OR 0.77, $p=0.4$) despite significantly lower odds of pulmonary
45 complications with pneumonectomy (matching OR 0.4, $p<0.0001$; weighting OR 0.12, $p<0.001$). The
46 adjusted hazard for death after pneumonectomy was significantly higher when analyzed using matched
47 analysis but not with weighting (Matching HR 1.63, $p=0.002$; weighting HR 0.97, $p=0.92$). The same
48 was true for disease-free survival (Matching HR 1.49, $p=0.01$; weighting HR 1.03, $p=0.84$).

49
50 **Conclusions:** Despite early differences in perioperative pulmonary outcomes favoring
51 pneumonectomy, early overall and disease-free survival were in favor of sleeve lobectomy in the
52 matched analysis but not the weighted analysis. In our opinion, when it is technically feasible, sleeve
53 lobectomy should be the preferred technique.

54

55 **Key-words:** Pneumonectomy, sleeve lobectomy, survival, outcomes

56 **Central message**

57 There were no differences in postoperative mortality between the two techniques. Sleeve lobectomy
58 increases the risk of pulmonary complications.

59

60 **Perspective statement**

61 Sleeve lobectomy leads to higher rate of overall and disease-free survival despite an increased risk of
62 postoperative pulmonary complications. We think that when it is technically possible, sleeve
63 lobectomy has to be the type of resection to favor for central tumor.

64

65 **Central Picture**

66 Weighting analysis did not highlight any benefit in overall survival for sleeve lobectomy.

67

68 **Abbreviations**

69 ARDS: Acute Respiratory Distress Syndrome

70 ASA: American Society of Anesthesiologists

71 BMI: Body Mass Index

72 BPF: Bronchopleural Fistula

73 CI: Confidence Interval

74 DFS: Disease-Free Survival

75 FEV1: Forced Expiratory Volume in one second

76 HR: Hazard Ratio

77 IPTW: Inverse Probability for Treatment Weighting

78 NSCLC: Non-small cell lung cancer

79 OR: Odds Ratio

80 OS: Overall Survival

81 PN: Pneumonectomy

82 PS: Propensity Score

83 SL: Sleeve Lobectomy

84 RCT: Randomized Control Trial

85 WHO: World Health Organization

86

87 Manuscript

88

89 Introduction

90 Complete surgical resection is the cornerstone of the management of localized non-small cell lung
91 cancer (NSCLC). For years, surgical resection was restricted to segmentectomy, lobectomy, and
92 pneumonectomy (PN), the last being associated with the highest postoperative morbidity and mortality
93 [1]. Sleeve Lobectomy (SL) was first described for carcinoma by Allison, in 1959, in order to avoid
94 the burden of PN in frail patients with compromised lung function [2]. Given the morbidity and
95 mortality associated with PN, sleeve lobectomy (SL) was initially restricted to patients with
96 compromised lung function who would not tolerate PN. However, the indication for SL has
97 progressively expanded to any tumor that may be completely resected using this technique,
98 particularly on the right side [3].

99 As a result of this progressive expansion of SL over PN, recent guidelines from the American College
100 of Chest Physicians recommend SL rather than PN in patients with clinical stage I or II central
101 NSCLC in whom complete resection can be achieved [4]. However, these guidelines are based on
102 short surgical series comparing recent SL to historical PN, or meta-analyses that included patients
103 operated on over many decades [5]. No database analysis has been published to date even though
104 every lung cancer surgeon is now able to perform SL and PN, and recent publications have highlighted
105 the impact of surgery improvements on the prognosis of NSCLC during recent decades [6]. We
106 therefore sought to compare short and long-term outcomes following SL and PN for NSCLC during
107 the last decade in France. For this purpose, we used the French national database Epithor and two PS
108 methods: matching and the inverse probability of treatment weighting (IPTW), combined with a
109 sensitivity analysis.

110

111 Materials and Methods

112 Data Collection

113 Epithor is a government-recognized clinical database, financially supported by the National Cancer
114 Institute (Institut National du Cancer) for data-quality monitoring. Epithor is accredited by the French

115 Health Authorities (Haute Autorité de Santé), a governmental agency dedicated to improving the
116 quality of patient care and to guaranteeing equality within the health care system, as a
117 methodologically appropriate tool to assess professional surgical practices. Participating in Epithor is
118 now a requirement for medical accreditation and thoracic surgery unit certification in France [7,8].

119 The accuracy of data collection is checked in regular external on-site audits initiated in 2010 [7]. Data
120 are sent by Internet to the national database; surgeons and patients are anonymous. Surgeons can
121 check the quality of the way they enter the data by comparing their data with national data through a
122 quality score ranging from 0% to 100%. Moreover, participants have to check the quality of the local
123 database for missing values by comparing its completeness with that of the national database. This
124 comparison is expressed through a quality score ranging from 0 to 100%. A score exceeding 80% is
125 mandatory to have the local data incorporated in the national database and to benefit from the
126 accreditation. Every surgeon receives a personal quality score, thus inciting them to update their data.
127 This induces a virtuous cycle. Almost all of the teams that participate in Epithor have a score above
128 80% for data entry.

129 The Institutional Review Board of the French Society of Thoracic and Cardiovascular Surgery
130 approved the electronic prospective database used for this study and the study itself. Patients' consent
131 was obtained, and patients were aware that the data collected would be used for clinical research
132 purposes.

133 The institutional review board of the French Society of Thoracic and Cardiovascular Surgery certifies
134 that this study respects the current regulations that govern clinical research in France, referenced as:
135 "CERC-SFCTCV-2015-8-14-16-5-39-PAPi.

136

137 ***Study population***

138 From January 2005 to December 2014, 6259 patients underwent SL or PN for NSCLC in 103 centers
139 in France. The baseline demographic and clinical characteristics included age, sex, American Society
140 of Anesthesiologists (ASA) score, World Health Organization (WHO) performance status, body mass
141 index (BMI), the Forced Expiratory Volume in 1 second (FEV1) as a percentage and the dyspnea
142 score according to the Medical Research Council [9]. The number of comorbid diseases per patient

143 was considered a categorical variable because recent data from Epithor consistently suggested that this
144 variable was superior to individual comorbidities in a predictive model for operative mortality [10].
145 Systematic nodal dissection included node sampling or radical lymphadenectomy. NSCLC histology
146 was classified according to the most recent WHO classification [11]. Tumor and Nodal stages were
147 classified postoperatively according to the pathologic examination and the most recent IASLC
148 classification [12].

149

150 *Outcome measurements*

151 The primary endpoint was postoperative complications, which included cardiopulmonary morbidity,
152 bronchopleural fistula (BPF), empyema, and hemorrhage. Cardiopulmonary morbidity was reported as
153 proposed by the European Society of Thoracic Surgery and included postoperative pulmonary
154 complications (pneumonia, atelectasis requiring bronchoscopy, adult respiratory distress syndrome
155 (ARDS), mechanical ventilation for more than 2 days and arrhythmia [13].

156 The secondary endpoints were (i) postoperative mortality, defined as any patient who died within the
157 first 30 days following surgery, or the initial hospitalization if longer; (ii) length of hospital stay, (iii)
158 overall survival (OS), defined as the time from surgery until death from any cause or the last follow-up
159 visit; and (iiii) disease-free survival (DFS), defined as the time from surgery until disease recurrence
160 or the last follow-up visit.

161

162 *Variables used for PS analysis*

163 Variables used to estimate the PS were age, sex, performance status, number of comorbidities,
164 dyspnea score, FEV1, BMI, induction chemotherapy, side, histology, T status, N status, year of
165 surgery, type of center and hospital volume.

166

167 *Missing data*

168 The proportion of missing FEV1 for this study was 7%. We assumed that the missing data were
169 missing at random. We applied a multiple imputation framework (20 imputations) to compensate for
170 missing prognostic factor data for FEV1.

171 For missing data for tumor pathological features, lymph nodes and histology, we created a variable
172 category to include in the analysis. For each hospital, the hospital volume was calculated for the period
173 from 2005 to 2014. Centers were ranked by number of procedures performed per year. We created a
174 categorical variable detailing five categories of hospital volume.

175

176 *Statistical analysis*

177 The PS is the conditional probability of assignment to a particular treatment given a vector of observed
178 covariates [14]. PS techniques were used to balance the distributions of measured potentially
179 confounding covariates for patients treated by SL or PN. A mirrored histogram was used to measure
180 the discriminatory ability of PS matching, and the standardized difference for the IPTW analysis.
181 Matching and IPTW tend to eliminate systematic differences between experimental and control
182 subjects to a greater degree than does stratification or covariate adjustment [15,16]. Matching used a
183 search algorithm to find a set of weights for each covariate such that the version of optimal balance is
184 achieved after matching [17]. We used 1:1 matching without replacement in descending order with a
185 caliper of 0.01. With IPTW, each individual is weighted by the inverse probability of receiving the
186 treatment that they actually received. In this way, each group is weighted to represent the full
187 population sample, thus revealing treatment effects. We evaluated two PS techniques in their ability to
188 balance the measured covariates between SL and PN by reducing the standardized difference [16]. The
189 standardized difference is the difference between sample means in the SL and PN group divided by the
190 standard deviation in the treatment group overall [17]. Finally, odds ratios (OR) were used for
191 dichotomous variables such as postoperative mortality, atelectasis, pneumonia, arrhythmia, BPF,
192 empyema and hemorrhage, ventilation > 2 days and ARDS. Logistic models were used. For the length
193 of hospital stay, the difference of means was used. Logistic models and linear regression were used.
194 For OS and DFS, we used the adjusted Kaplan-Meier estimators for curves using IPTW data and the
195 log-rank test. Cox proportional hazards models were used to estimate Hazard ratios (HR). Hospital
196 level clustering was used to estimate the robust standard error for each coefficient of the logistic
197 model, linear regression and Cox model.

198

199 *Sensitivity analysis*

200 The Q Mantel-Haenszel test statistic was used; if $\hat{\Gamma}=1$, the statistic test is significant and the study is
201 free of hidden bias [19]. And if the value of the Q Mantel-Haenszel test is greater, then the study is
202 insensitive to bias [20].

203

204 **Results**

205 *Study cohort*

206 From 2005 to 2014, 941 SL and 5318 PN were included in the Epithor database and make up the study
207 cohort. As compared to the PN group, patients in the SL group were characterized by a younger age
208 (median 62 years old for both groups), higher BMI (median 25 for both groups), higher FEV1, lower
209 ASA score, lower WHO status, and less frequent induction therapy (**Table 1**). As compared to the PN
210 group, the SL group was characterized by the predominance of right-sided surgery and squamous cell
211 carcinoma. T and N stages in the SL group were lower than those in the PN group (**Table 1**).

212

213 *Surgery*

214 More than half of the SL involved the right upper lobe, followed by the left upper, left lower and right
215 lower respectively. Middle lobectomy and bilobectomy SL occurred in less than 5% of patients each
216 (Table 1). Both SL and PN were predominantly performed in teaching hospitals. SL was more
217 frequently performed in high-volume centers, whereas PN was more frequently performed in low-
218 volume centers. The number of SL increased in the last decade whereas the number of PN decreased
219 steadily (**Table 1**).

220

221 *PS estimation*

222 The mirrored histogram shows the good distribution of the covariates after PS matching (Figure 1).
223 The median distribution of standardized biases was 0.024 for the matching approach (1st and 3rd
224 quartile: 0.013-0.0355), and 0.04 for IPTW (1st and 3rd quartile: 0.0145-0.0654). The standardized
225 difference never reached the value of 10%, which highlighted that the two groups were well balanced
226 for covariates by matching and IPTW [19] (Table 2).

227

228 Postoperative Mortality

229 Postoperative mortality was 4.99% (n=47) in the SL group and 5.89% (n=313) in the PN group
230 (p=0.279). There were no significant differences in post-operative mortality between SL and PN,
231 according to the 2 PS methods, with an OR associated with PN of 1.24 (95%CI) [0.74-2.1] for
232 matching, and 0.77 (95%CI) [0.4-1.5] for IPTW (OR=1 for SL, **Table 3**).

233

234 Postoperative complications and Length of Hospital stay

235 Postoperative pulmonary complications overall (pneumonia, atelectasis requiring bronchoscopy,
236 ARDS, mechanical ventilation for more than 2 days) were significantly less frequent in the PN group
237 than in the SL group according to the two PS methods (**Table 3**). In detail, only atelectasis and
238 pneumonia were significantly less frequent following PN whatever the PS analysis performed. The
239 two PS methods showed a trend towards a higher rate of postoperative arrhythmia, and a significantly
240 higher rate of BPF and empyema in the PN group than in the SL group (**Table 3**). The incidence of
241 hemorrhage was significantly higher in the PN group but only by matching (**Table 3**). BPF occurred in
242 2.6% of patients in the PN group (n=138) and 1.59% in SL group (n=15) (p=0.067). In the PN group,
243 there was no significant difference in the occurrence of BPF whether the patient receive preoperative
244 irradiation (3.17%, n=4) or not (2.34%, n=122) (p=0.679). Length of hospital stay was significantly
245 shorter in the SL group than in the PN group by matching but not by IPTW analysis (**Table 3**).

246

247 Overall Survival and Disease-free Survival

248 The median follow-up time was 10.89 months (1st and 3rd quartile: 1.66-15) for OS and 9.6 months
249 (1st and 3rd quartile: 1-14) for DFS. From the 941 patients of the SL group, 463 were evaluable at 1
250 month, 174 at 12 months, 39 at 36 months and 10 at 60 months. From the 5318 patients of the PN
251 group, 2222 were evaluable at 1 month, 719 at 12 months, 154 at 36 months and 20 at 60 months.

252 Three-year OS was $71.86 \pm 3.29\%$ for the SL group and $60.76 \pm 1.69\%$ for the PN group. As
253 compared with SL, there was an increased risk of death in the PN group by matching, with HR of 1.63
254 [1.19-2.21], but not by the IPTW method (Table 3, Figure 2).

255 Three-year DFS was $46.41 \pm 4.01\%$ for the SL group and $31.63 \pm 1.59\%$ for the PN group. As
256 compared with SL, there was an increased risk of recurrence in the PN group by matching, with HR of
257 1.49 [1.1-2], but not by the IPTW method (Table 3 & Figure 3).

258

259 *Sensitivity analysis*

260 For the matching analysis, the Q Mantel-Haenszel (Qmh) statistic test was used and showed a
261 potential hidden bias (Qmh value close to 1) only for postoperative mortality (**Supplementary Table**
262 **4**). For postoperative complications, the Qmh value was high, indicating the study was insensitive to a
263 bias that would the odds [19]. These results confirmed that compared to SL, PN was associated with a
264 significant decrease in the rate of pulmonary complications overall (pneumonia, atelectasis requiring
265 bronchoscopy, ARDS, mechanical ventilation for more than 2 days), a significant decrease in the rate
266 of atelectasis and pneumonia considered separately, a non-significant increase in the rate of
267 arrhythmias, and a significant increase in the risk of BPF, empyema and hemorrhage. For the IPTW
268 method, the sensitivity analysis confirmed these data, except for arrhythmias and hemorrhages
269 (**Supplementary Table 5**).

270

271 **Discussion**

272 *Reminder of the main results*

273 In studying outcomes after PN and SL for NSCLC using data from a nationwide database, two PS
274 methods concluded that SL was not associated with any significant difference in postoperative
275 mortality, but with a significant increase in the rate of pulmonary complications (atelectasis,
276 pneumonia), and a significant decrease in the rate of bronchopleural fistulae and empyema as
277 compared with PN. Matching analysis also found that SL was associated with improved 3-year OS as
278 compared with PN, but not IPTW analysis.

279

280 *Treatment allocation*

281 Initially developed for patients with insufficient pulmonary reserve, which did not allow PN, SL has
282 now become widely accepted as a reliable and safe procedure to allow the complete resection of

283 NSCLC [22]. Indeed, SL makes it possible to spare lung parenchyma, and thus provides better
284 postoperative FEV1 than that achieved with PN [23–25]. Quality of life after SL is also better than that
285 after PN, as highlighted by Balduyck et al., who reported a high burden of dyspnea, general pain,
286 thoracic pain and shoulder dysfunction after PN [26]. This is the first time, however, that the impact of
287 SL on postoperative complications, mortality, and long-term survival had been reported in the context
288 of a large study based on a national database, which allows the analysis of real-life events. In this
289 respect, one can assume that over the last decade all French thoracic surgeons were able to perform
290 both procedures. The choice to perform one procedure rather than the other was therefore based on
291 disease severity and the patient's clinical status rather than technical preferences. This probable
292 difference between groups in baseline characteristics of patients and tumors constitutes the strongest
293 argument in favor of adequate statistical analysis, including PS analysis, of this large real-life dataset.

294

295 ***Postoperative mortality and postoperative complications***

296 We found no significant difference regarding postoperative mortality between SL and PN groups,
297 whatever the statistical method used. In a recent paper from the Epithor Group, Morgant et al. reported
298 PN and bronchial SL to be major prognostic factors associated with postoperative death following
299 lung cancer surgery, with OR ranging from 4.4 to 8.2 for PN and 2.4 to 2.9 for bronchial SL [6]. In a
300 recent meta-analysis that included more than 3800 patients, Shi et al. showed a significant difference
301 in postoperative mortality favoring SL over PN with an OR of 0.5 [0.34-0.72] [5]. Similarly, in a
302 matching analysis comparing SL and PN, Park et al. found postoperative mortality to be significantly
303 lower in the SL group (1%) than in the PN group (8.6%, $p < 10^{-4}$) [27]. In contrast, we found no
304 difference between SL and PN for postoperative mortality, despite a significantly greater incidence of
305 BPF and empyema after PN. We found no impact of preoperative irradiation on the occurrence of BPF
306 after PN, as already highlighted in a recent paper establishing a predictive score for BPF [28]. This
307 absence of difference can be attributed either to the low mortality associated with PN, or to the high
308 mortality associated with SL in our study.

309 On the one hand, recent publications tend to disfavor PN, which, for example, was associated with a
310 mortality rate of 7.8% over the last decade in the French database [28]. On the other hand, in our

311 study, mortality following SL remained far higher than that reported in the most recent studies
312 [5,22,24], and markedly higher than the 2.7% reported after regular lobectomy in the same database
313 [7]. Indeed, in our study, postoperative atelectasis and pneumonia were significantly more frequent
314 after SL than after PN, whatever the statistical method considered. This high rate of parenchymal
315 complications may explain the increased postoperative mortality after SL than after regular lobectomy
316 [7]. Altogether, these data suggest that the improvement in patient selection and postoperative care
317 developed in PN should be applied to SL in order to decrease its postoperative morbidity and
318 mortality.

319

320 ***Overall Survival and Disease-free Survival***

321 We found that OS was lower in the PN group than in the SL group only by matching. Similar results
322 were reported in the meta-analysis by Shi et al, with a combined HR of 0.63 [0.56-0.71], and in the
323 matching analysis by Park et al. (5-year OS for PN vs SL, 32.1% vs 58.4%, respectively, $p=0.0002$)
324 [5,27]. In our study, patients undergoing PN were significantly more likely to have T3 and T4 tumors,
325 and N2 lymph-node involvement than were patients undergoing SL. The TNM classification alone
326 might lack precision: T stages have recently been separated according to tumor size [29], the N stage
327 might need to be deciphered according to the number of lymphatic chains and stations involved [30],
328 and the number of molecular biomarkers is constantly increasing [31]. However, as in our study, the
329 meta-analysis of Shi et al. showed significantly more advanced disease in the PN group (47.96% of
330 Stage III in the PN group vs. 38.32% in the SL group; $p<0.001$), which could explain by itself the
331 worse prognosis in patients who underwent pneumonectomy [5].

332 Therefore, PS methods tend to balance the distributions of potentially confounding covariates, but the
333 lack of precision of the covariates measured might limit the impact of PS in this setting. As compared
334 with the SL group, the PN group had significantly more recurrences according to matching. These
335 findings were different from those of Park et al. whose sample size was too small to draw any
336 definitive conclusions [27], or by Shi et al., whose matching analysis allowed a more powerful
337 comparison of OS and DFS following SL and PN [5].

338

339 However, the two PS analyses did not provide the same results, and the lower OS and DFS in the PN
340 group were not significant in the IPTW analysis. It is important to bear in mind that, unlike
341 randomized assignments of treatments, PS typically does not balance covariates that were not
342 observed [32]. Matching reduces the sample size since matches may not be found for some patients,
343 whereas weighting allows the comparison of expectations and distributions between treated and
344 untreated subjects [16,33]. The combination of matching and IPTW tends to eliminate systematic
345 differences between experimental and control subjects to a greater degree than does stratification or
346 covariate adjustment [15,16]. Even if these two techniques well balanced the distribution of the
347 covariates, they do not permit to conclude which one provide the “true results” and also raised the
348 question of the reality of the results published in studies using only matching analysis. Given the
349 differences in the results of matching and IPTW, the results for long-term survival and recurrence
350 should be interpreted with caution.

351

352 *Strengths and limitations*

353 To the best of our knowledge, this is the first study to compare SL and PN using a large dataset, a
354 national database, two PS methods and a sensitivity analysis. The main strengths of this study are the
355 use of a national database, which provided a large number of patients, a homogeneous population. The
356 large number of patients in both groups allowed powerful comparisons. The homogeneous population
357 reduced the sample size needed for the matching comparison and the reasonable length of study period
358 tended to decrease historical bias.

359 However, any study involving a large database raises the question of the quality and exhaustiveness of
360 the prospectively entered data, such as comorbidities, and observational studies are notoriously full of
361 no responses and missing values [28]. Few details were available about the surgical technique,
362 especially concerning bronchial stump coverage and pulmonary artery reconstruction. We used PS to
363 create comparable cohorts; however, we cannot be certain that the PS perfectly neutralized all of the
364 confounding variables, as suggested by the differences in the results for the two methods used.

365 The main limitation of this study is the dramatic loss of follow-up data (50% of missing data after only
366 1 month of follow-up), which could be explained by various factors: at the visit 1 month after the

367 surgery, some patients are not seen by the thoracic surgeon but by the medical oncologist or the
368 pneumologist; and some of the surgeons do not always enter follow-up data into the Epithor database,
369 for example, when they hear news about the patients or correspondence announcing their death.
370 Moreover, we must underline that PS cannot replace randomized control trials (RCTs). In RCTs,
371 random allocation of patients to either an experimental or a control arm guarantees that treatment
372 allocation is unrelated to measured and unmeasured patients' characteristics. It enables researchers to
373 draw unbiased conclusions about a treatment effect, provided that the number of randomized patients
374 is large enough to minimize random variation [33]. Even though this study has a high level of
375 evidence, the conclusions that can be drawn will never be as strong as those from prospective
376 multicenter RCT, which seem to be impossible to conduct nowadays.

377

378 **Conclusion**

379 Despite the limits of our study, especially the dramatic loss of follow-up data, we highlighted the
380 interest of SL, which could lead to better 3-year OS and DFS as compared to PN for NSCLC patients.
381 Whenever it is technically possible, surgeons must perform SL in order to provide more long-term
382 survival benefits to patients even with the risk of more postoperative pulmonary complications.

383

384 **Acknowledgments**

385 This work was performed on behalf of the French Society of Thoracic and Cardiovascular surgery
386 (Société Française de Chirurgie Thoracique et CardioVasculaire – SFCTCV). We are grateful to all
387 the French thoracic surgeons who participated in this study and have listed them in e-appendix 1.
388 Pierre-Benoit Pagès is the recipient of research grants allocated by the Nuovo-Soldati foundation for
389 Cancer Research (Fondation Nuovo-Soldati) and the French Society of Thoracic and Cardiovascular
390 Surgery (SFCTCV).

391 The authors would like to thank Philip Bastable (Dijon University Hospital) for reviewing the
392 manuscript.

393

394 **References**

395

396 [1] Fell SC. A history of pneumonectomy. *Chest Surg Clin N Am* 1999;9:267–290, ix.

397 [2] Allison PR. Course of thoracic surgery in Groningen. Cited by Jones PV. *Ann R Coll Surg*
398 *Engl* 1959;25:30-38.

399 [3] Deslauriers J, Grégoire J, Jacques LF, Piraux M, Guojin L, Lacasse Y. Sleeve lobectomy
400 versus pneumonectomy for lung cancer: a comparative analysis of survival and sites of recurrences.
401 *Ann Thorac Surg* 2004;77:1152–1156; discussion 1156.

402 [4] Howington JA, Blum MG, Chang AC, Balekian AA, Murthy SC. Treatment of stage I and II
403 non-small cell lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of
404 Chest Physicians evidence-based clinical practice guidelines. *Chest* 2013;143:e278S–313S.

405 [5] Shi W, Zhang W, Sun H, Shao Y. Sleeve lobectomy versus pneumonectomy for non-small cell
406 lung cancer: a meta-analysis. *World J Surg Oncol* 2012;10:265.

407 [6] Morgant M-C, Pagès P-B, Orsini B, Falcoz P-E, Thomas P-A, Barthes FLP, et al. Time trends
408 in surgery for lung cancer in France from 2005 to 2012: a nationwide study. *Eur Respir J* 2015.

409 [7] Thomas PA, Berbis J, Falcoz P-E, Le Pimpec-Barthes F, Bernard A, Jougon J, et al. National
410 perioperative outcomes of pulmonary lobectomy for cancer: the influence of nutritional status. *Eur J*
411 *Cardiothorac Surg* 2014;45:652–659; discussion 659.

412 [8] Pagès P-B, Delpy J-P, Falcoz P-E, Thomas P-A, Filaire M, Le Pimpec Barthes F, et al.
413 Videothoracoscopy versus thoracotomy for the treatment of spontaneous pneumothorax: a propensity
414 score analysis. *Ann Thorac Surg* 2015;99:258–63.

415 [9] C.S. Dyspnea score The MRC breathlessness scale. (Lond). COM, editor, 2008:226-7.

416 [10] Bernard A, Rivera C, Pages PB, Falcoz PE, Vicaut E, Dahan M. Risk model of in-hospital
417 mortality after pulmonary resection for cancer: a national database of the French Society of Thoracic
418 and Cardiovascular Surgery (Epithor). *J Thorac Cardiovasc Surg* 2011;141:449–58.

419 [11] Travis WD, Brambilla E, Müller-Hermelink HK et al. Pathology and Genetics: Tumours of the
420 Lung, Pleura, Thymus and Heart. Lyon: IARC, 2004.

421 [12] Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, et al. The IASLC
422 Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the

- 423 forthcoming (seventh) edition of the TNM Classification of malignant tumours. *J Thorac Oncol*
424 2007;2:706–14.
- 425 [13] Brunelli A, Berrisford RG, Rocco G, Varela G, European Society of Thoracic Surgeons
426 Database Committee. The European Thoracic Database project: composite performance score to
427 measure quality of care after major lung resection. *Eur J Cardiothorac Surg* 2009;35:769–74.
- 428 [14] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies
429 for causal effects. *Biometrika* 1983;70:41-55.
- 430 [15] Austin PC, Grootendorst P, Anderson GM. A comparison of the ability of different propensity
431 score models to balance measured variables between treated and untreated subjects: a Monte Carlo
432 study. *Stat Med* 2007;26:734–53.
- 433 [16] Austin PC. The relative ability of different propensity score methods to balance measured
434 covariates between treated and untreated subjects in observational studies. *Med Decis Making*
435 2009;29:661–77.
- 436 [17] Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of
437 measured covariate balance to test causal associations in psychological research. *Psychol Methods*
438 2010;15:234–49.
- 439 [18] Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci*
440 2010;25:1–21. doi:10.1214/09-STS313.
- 441 [19] Mantel, N., and W. Haenszel. 1959. Statistical aspects of the analysis of data from
442 retrospective studies. *Journal of the National Cancer Institute* 22: 719–748.
- 443 [20] Rosenbaum PR 2002. *Observational studies* 2nd ed New York: Springer.
- 444 [21] Normand ST, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, et al. Validating
445 recommendations for coronary angiography following acute myocardial infarction in the elderly: a
446 matched analysis using propensity scores. *J Clin Epidemiol* 2001;54:387–98.
- 447 [22] Nagayasu T, Yamasaki N, Tsuchiya T, Matsumoto K, Miyazaki T, Hatachi G, et al. The
448 evolution of bronchoplasty and broncho-angioplasty as treatments for lung cancer: evaluation of 30
449 years of data from a single institution. *Eur J Cardiothorac Surg* 2015.

- 450 [23] Martin-Ucar AE, Chaudhuri N, Edwards JG, Waller DA. Can pneumonectomy for non-small
451 cell lung cancer be avoided? An audit of parenchymal sparing lung surgery. *Eur J Cardiothorac Surg*
452 2002;21:601–5.
- 453 [24] Melloul E, Egger B, Krueger T, Cheng C, Mithieux F, Ruffieux C, et al. Mortality,
454 complications and loss of pulmonary function after pneumonectomy vs. sleeve lobectomy in patients
455 younger and older than 70 years. *Interact Cardiovasc Thorac Surg* 2008;7:986–9.
- 456 [25] Gómez-Caro A, Boada M, Reguart N, Viñolas N, Casas F, Molins L. Sleeve lobectomy after
457 induction chemoradiotherapy. *Eur J Cardiothorac Surg* 2012;41:1052–8.
- 458 [26] Balduyck B, Hendriks J, Lauwers P, Van Schil P. Quality of life after lung cancer surgery: a
459 prospective pilot study comparing bronchial sleeve lobectomy with pneumonectomy. *J Thorac Oncol*
460 2008;3:604–8.
- 461 [27] Park JS, Yang HC, Kim HK, Kim K, Shim YM, Choi YS, et al. Sleeve lobectomy as an
462 alternative procedure to pneumonectomy for non-small cell lung cancer. *J Thorac Oncol* 2010;5:517–
463 20.
- 464 [28] Pforr A, Pagès PB, Baste JM, Falcoz PE, Thomas P, Le Pimpec-Barthes F et al.
465 Bronchopleural fistula predictive score: setting up from the French Database Epithor. *Ann Thorac*
466 *Surg.* 2016 Jan;101(1):287-93.
- 467 [29] Thomas PA, Berbis J, Baste J-M, Le Pimpec-Barthes F, Tronc F, Falcoz P-E, et al.
468 Pneumonectomy for lung cancer: contemporary national early morbidity and mortality outcomes. *J*
469 *Thorac Cardiovasc Surg* 2015;149:73–82.
- 470 [30] Rami-Porta R, Bolejack V, Crowley J, Ball D, Kim J, Lyons G, et al. The IASLC Lung Cancer
471 Staging Project: Proposals for the Revisions of the T Descriptors in the Forthcoming Eighth Edition of
472 the TNM Classification for Lung Cancer. *J Thorac Oncol* 2015;10:990–1003.
- 473 [31] Legras A, Mordant P, Arame A, Foucault C, Dujon A, Le Pimpec Barthes F, et al. Long-term
474 survival of patients with pN2 lung cancer according to the pattern of lymphatic spread. *Ann Thorac*
475 *Surg* 2014;97:1156–62.

- 476 [32] Rami-Porta R, Asamura H, Goldstraw P. Predicting the prognosis of lung cancer: the
477 evolution of tumor, node and metastasis in the molecular age-challenges and opportunities. *Transl*
478 *Lung Cancer Res* 2015;4:415–23.
- 479 [33] Joffe MM, Rosenbaum PR. Invited commentary: propensity scores. *Am J Epidemiol*
480 1999;150:327–33.
- 481 [34] Heinze G, Jüni P. An overview of the objectives of and the approaches to propensity score
482 analyses. *Eur Heart J* 2011;32:1704–8.
- 483

484 **Table 1**

485 Characteristics of patients undergoing sleeve lobectomy and pneumonectomy.

		<u>Full Sample</u>		P Value
		Sleeve Lobectomy Group (n=941)	Pneumonectomy Group (n=5318)	
<u>Demographics</u>				
Sex	Male	716 (76.1%)	4216 (79.3%)	0.027
	Female	225 (23.9%)	1102 (20.7%)	
Age (Years)		60.9 ± 12.6	61.9 ± 10.2	0.014
Body Mass Index (Kg/m²)		25.6 ± 4.5	25.1 ± 4.1	0.0009
ASA	1	181 (19.2%)	817 (15.3%)	0.006
	2	522 (54.5%)	2951 (55.5%)	
	3	224 (23.8%)	1434 (27%)	
	4	14 (1.5%)	116 (2.2%)	
WHO Performance Status	0	442 (47%)	2124 (40%)	<0.0001
	1	420 (44.6%)	2574 (48.4%)	
	2	72 (7.6%)	523 (9.8%)	
	3	7 (0.8%)	97 (1.8%)	
FEV (%)		74.1 ± 17.6	62.9 ± 20.5	<0.00001
Dyspnea score	0	434 (46.1%)	2596 (48.8%)	0.37
	1	328 (34.9%)	1831 (34.4%)	
	2	148 (15.8%)	715 (13.5%)	
	3	21 (2.2%)	121 (2.3%)	
	4	10 (1%)	55 (1%)	
Number of comorbidities per patient	0	131 (13.9%)	702 (13.2%)	0.003
	1	226 (24%)	1471 (27.6%)	
	2	271 (28.8%)	1543 (29%)	
	3	252 (26.8%)	1393 (26.2%)	
	4	61 (6.5%)	209 (4%)	
<u>Perioperative management</u>				
Neoadjuvant Chemotherapy		167 (17.7%)	1282 (24.2%)	<0.0001
Neoadjuvant Radiotherapy		11 (1.2%)	126 (2.4%)	0.02
Adjuvant Chemotherapy for N1		202 (78%)	2378 (64%)	0.466
N2		135 (90%)	1138 (83%)	0.4
<u>Tumor characteristics</u>				
Side	Right	652 (69.6%)	2172 (41%)	<0.0001
	Left	285 (30.4%)	3122 (59%)	
Tumor	T1	190 (20.2%)	347 (6.5%)	<0.0001
	T2	408 (43.3%)	1813 (34.1%)	
	T3	169 (18%)	1482 (27.9%)	
	T4	32 (3.4%)	690 (13%)	
	Missing	142 (15.1%)	986 (18.5%)	
Lymph nodes	N0	385 (40.9%)	1394 (26.2%)	<0.0001
	N1	259 (27.5%)	1522 (28.6%)	

	N2	150 (16%)	1372 (25.8%)	
	N3	1 (0.1%)	22 (0.4%)	
	Missing	146 (15.5%)	1008 (19%)	
Postoperative Histology	Squamous	514 (54.6%)	2570 (48.3%)	<0.0001
	Adenocarcinoma	160 (17%)	1602 (30.1%)	
	Carcinoid	138 (14.7%)	69 (1.3%)	
	Others	57 (6%)	477 (9%)	
	Missing	72 (7.7%)	600 (11.3%)	
Resection Margins R0		797 (84.7%)	4388 (82.5%)	<0.0001
	R1	66 (6.4%)	232 (4.4%)	
	R2	3 (0.3%)	47 (0.9%)	
	Missing	81 (8.6%)	651 (12.2%)	
<u>Characteristics of Center</u>				
Type of center	Non-Academic	53 (5.6%)	669 (12.6%)	<0.0001
	Private	265 (28.2%)	1556 (29.3%)	
	Academic	623 (66.2%)	3087 (58.1%)	
Hospital volume of activity (number of procedures/year)	<48	136 (14.5%)	1235 (23.2%)	<0.0001
	49-84	193 (20.5%)	1023 (19.2%)	
	85-133	163 (17.3%)	903 (17%)	
	134-171	213 (22.6%)	1214 (22.8%)	
	>171	236 (25.1%)	943 (17.8%)	
Year of treatment	2005-2006	175 (18.6%)	1003 (18.9%)	0.005
	2007-2008	164 (17.4%)	1189 (22.4%)	
	2009-2010	207 (22%)	1174 (22%)	
	2011-2012	189 (20%)	938 (17.6%)	
	2013-2014	206 (22%)	1014 (19.1%)	

486

487 ASA: American Society of Anesthesiologists; FEV: Forced Expiratory Volume; WHO: World Health

488 Organization

489

490 **Table 2**

491 Patients' baseline characteristics for sleeve lobectomy and pneumonectomy (in parentheses percentage or mean) with their standardized difference. A
 492 standardized difference greater than 0.1 (10%) represents meaningful imbalance in a given variable between treatment groups.

Variables	Full sample			Matched			Weighted		
	Sleeve (n=941)	PN (n=5318)	standardized difference	Sleeve (n=794)	PN (n=794)	standardized difference	Sleeve (n=941)	PN (n=5318)	standardized difference
<u>Demographics</u>									
Sex Male	715 (76%)	4201 (79%)	0.078	635 (80%)	627 (79%)	-0.015	734 (78%)	4201 (79%)	0.017
BMI	25	25.6	-0.107	25.5	25.5	0.001	25.3	25.2	-0.028
WHO status									
1	423 (45%)	2552 (48%)	0.075	381 (48%)	381 (48%)	0.013	489 (52%)	2552 (48%)	-0.064
2	75 (8%)	532 (10%)	0.081	63 (8%)	63 (8%)	0.00001	84 (9%)	532 (10%)	0.006
3	9 (1%)	106 (2%)	0.094	8 (1%)	8 (1%)	-0.0022	28 (3%)	106 (2%)	-0.082
FEV	74%	63%	-0.583	73%	73%	-0.011	66.5%	64.7%	-0.098
Dyspnea score									
1	329 (35%)	1808 (34%)	-0.009	119 (15%)	301 (38%)	0.029	320 (34%)	1861 (35%)	0.009
2	150 (16%)	691 (13%)	-0.067	119 (15%)	119 (15%)	-0.021	132 (14%)	744 (14%)	-0.001
3	18 (2%)	106 (2%)	0.003	24 (3%)	16 (2%)	-0.034	37 (4%)	53 (1%)	-0.124
4	9 (1%)	53 (1%)	-0.003	8 (1%)	8 (1%)	-0.012	19 (2%)	53 (1%)	-0.083
Number of comorbidities per patients									
1	226 (24%)	1489 (28%)	0.083	190 (24%)	182 (23%)	-0.02	263 (28%)	1435 (27%)	-0.017
2	273 (29%)	1542 (29%)	0.003	238 (30%)	246 (31%)	0.011	244 (26%)	1542 (29%)	0.065
3	254 (27%)	1382 (26%)	-0.01	222 (28%)	222 (28%)	0.014	291 (31%)	1382 (26%)	-0.112
4	66 (7%)	212 (4%)	-0.115	47 (6%)	47 (6%)	-0.04	47 (5%)	212 (4%)	-0.01
<u>Perioperative management</u>									
Neoadjuvant chemotherapy	169 (18%)	1276 (24%)	0.156	159 (20%)	167 (21%)	0.009	216 (23%)	1276 (24%)	0.013

<u>Tumor characteristics</u>									
Right side	659 (70%)	2180 (41%)	-0.599	532 (67%)	516 (65%)	-0.042	423 (45%)	2393 (45%)	0.008
Tumor T2	404 (43%)	1808 (34%)	-0.19	349 (44%)	333 (42%)	-0.047	3667 (39%)	1914 (36%)	-0.062
T3	169 (18%)	1489 (28%)	0.241	166 (21%)	174 (22%)	0.036	216 (23%)	1382(26%)	0.088
T4	28 (3%)	691 (13%)	0.354	32 (4%)	47 (6%)	0.075	84 (9%)	638 (12%)	0.081
Missing	141 (15%)	957 (18%)	0.088	135 (17%)	127 (16%)	-0.017	188 (20%)	957 (18%)	-0.049
Lymph Nodes N1	263 (28%)	1542 (29%)	0.025	238 (30%)	230 (29%)	-0.036	254 (27%)	1542 (29%)	0.037
N2	150 (16%)	1382 (26%)	0.249	143 (18%)	150 (19%)	0.028	225 (24%)	1329 (25%)	0.026
N3	0 (0%)	0 (0%)	0.061	0 (0%)	8 (1%)	0.074	0 (0%)	0 (0%)	0.03
Missing	150 (16%)	1010 (19%)	0.087	135 (17%)	135 (17%)	-0.02	197 (21%)	1010 (19%)	-0.068
<u>Postoperative histology</u>									
Adenocarcinoma	160 (17%)	1595 (30%)	0.317	159 (20%)	159 (20%)	0.003	244 (26%)	1489 (28%)	0.054
Carcinoid	141 (15%)	53 (1%)	-0.51	40 (5%)	47 (6%)	0.019	28 (3%)	159 (3%)	-0.02
Others	56 (6%)	478 (9%)	0.111	55 (7%)	63 (8%)	0.038	94 (10%)	478 (9%)	-0.043
Missing	75 (8%)	584 (11%)	0.121	71 (9%)	63 (8%)	-0.026	94 (10%)	585 (11%)	0.01
<u>Characteristics of Center</u>									
Type of center Private	263 (28%)	1542 (29%)	0.02	230 (29%)	214 (27%)	-0.031	244 (26%)	1542 (29%)	-0.022
Academic	621 (66%)	3084 (58%)	-0.164	516 (65%)	508 (64%)	-0.013	564 (60%)	3137 (59%)	-0.022
<u>Hospital volume of activity (number of procedures/year)</u>									
49-84	197 (21%)	1010 (19%)	-0.032	159 (20%)	150 (19%)	-0.025	179 (19%)	1010 (19%)	-0.005
85-133	160 (17%)	904 (17%)	-0.008	135 (17%)	143 (18%)	0.027	179 (19%)	904 (17%)	-0.058
134-171	207 (22%)	1223 (23%)	0.012	190 (24%)	190 (24%)	-0.00001	216 (23%)	1223 (23%)	0.007
>171	235 (25%)	957 (18%)	-0.18	182 (23%)	174 (22%)	-0.028	169 (18%)	1010 (19%)	0.016
Year 2007-2008	169 (18%)	1169 (22%)	0.122	150 (19%)	135 (17%)	-0.035	225 (24%)	1169 (22%)	-0.05
2009-2010	207 (22%)	1169 (22%)	0.005	174 (22%)	182 (23%)	0.04	225 (24%)	1169 (22%)	0.046
2011-2012	188 (20%)	957 (18%)	-0.062	159 (20%)	174 (22%)	0.058	150 (16%)	957 (18%)	0.05
2013-2014	207 (22%)	1010 (19%)	-0.071	159 (20%)	150 (19%)	-0.031	188 (20%)	1010 (19%)	-0.01
<u>Overall propensity score</u>									
Mean	0.134			0.0258			0.040		
Median	0.085			0.0255			0.029		

493 **Table 3**

494 Effects of pneumonectomy vs sleeve lobectomy according to different statistical analyses (OR for
 495 SL=1, if OR <1 means a protective effect of pneumonectomy, OR>1 means a deleterious effect of
 496 pneumonectomy).

	Matching 1:1		Weighting	
		P value		P value
Postoperative mortality*	1.24 (0.74 – 2.1)	0.4	0.77 (0.4-1.5)	0.4
Postoperative complications*				
Postoperative pulmonary complications	0.4 (0.3-0.6)	<0.0001	0.12 (0.08-0.2)	<0.0001
Atelectasis	0.14 (0.08-0.25)	<0.0001	0.56 (0.36-0.87)	<0.0001
Pneumonia	0.4 (0.24-0.69)	<0.001	0.43 (0.31-0.61)	0.01
Ventilation > 2 days	1.4 (0.8-2.5)	0.2	1.25 (0.7-2.3)	0.4
ARDS	1.4 (0.7-2.5)	0.3	1.22 (0.6-2.5)	0.5
Arrhythmia	1.6 (0.95-2.6)	0.08	1.67 (0.9-2.9)	0.06
Bronchopleural fistula	2.9 (1.4-6.2)	0.005	2.5 (1.3-4.7)	0.004
Empyema	7 (1.9-27)	0.004	15 (4-57)	<0.0001
Hemorrhage	4 (1.9-8)	<0.0001	2 (0.7-6)	0.2
Length of hospital stay**	-1.31 (-2.54--0.09)	0.03	-1.18 (-3-0.63)	0.2
Overall survival***	1.63 (1.19-2.21)	0.002	0.97 (0.63-1.51)	0.92
Disease-Free survival***	1.49 (1.1-2.01)	0.01	1.03 (0.73-1.45)	0.84

497

498 *Odds Ratio (OR) with 95% confidence interval

499 **Difference of the mean number of days

500 ***Hazard ratio with 95% confidence interval

501 ARDS: Acute Respiratory Distress Syndrome

502

503

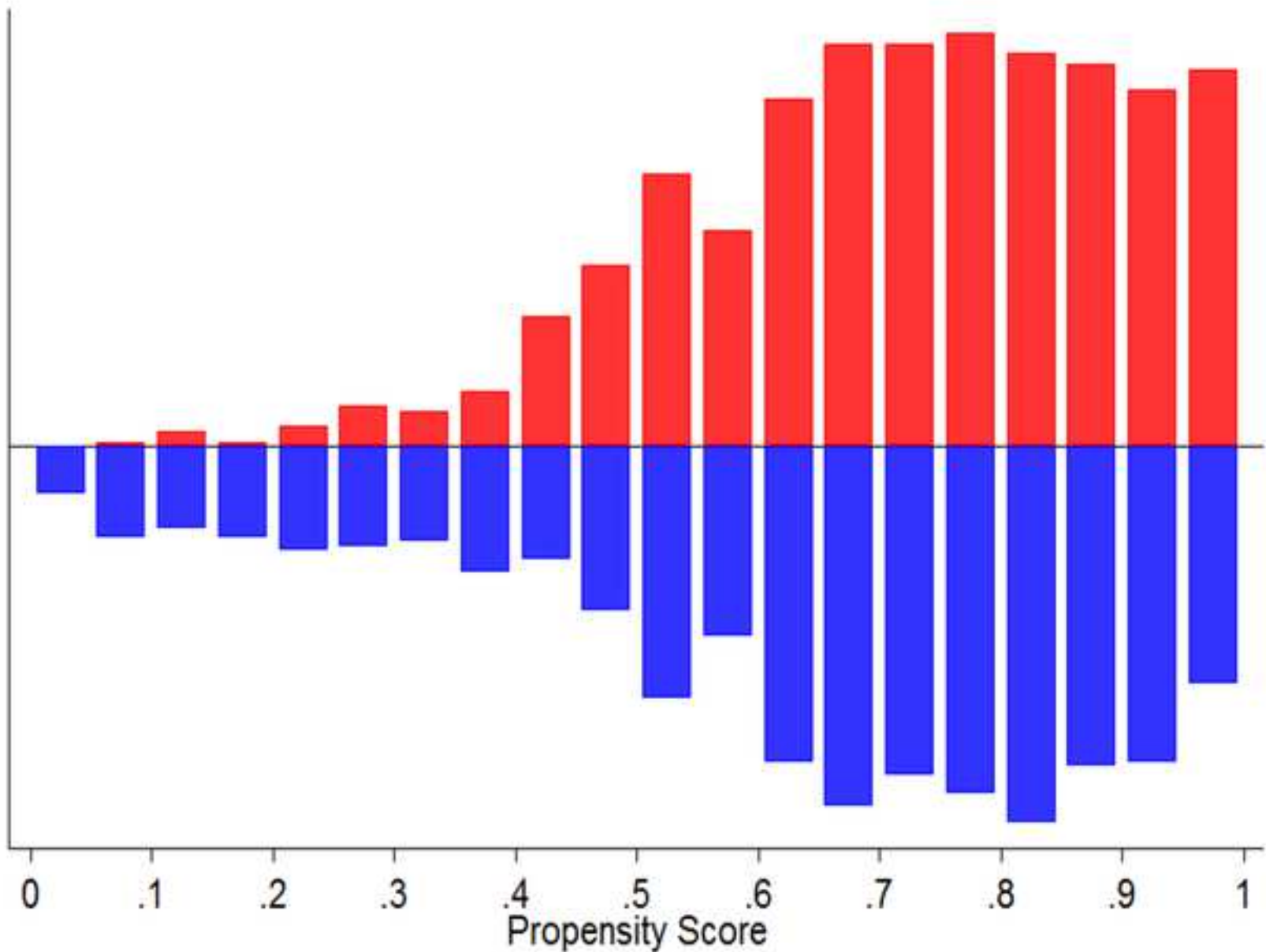
Figure legends



Figure 1: Kaplan-Meier plots of overall survival from the unmatched full sample (A), matched (B) and weighted (C) data for PN (continuous red line) and SL (discontinuous blue line).

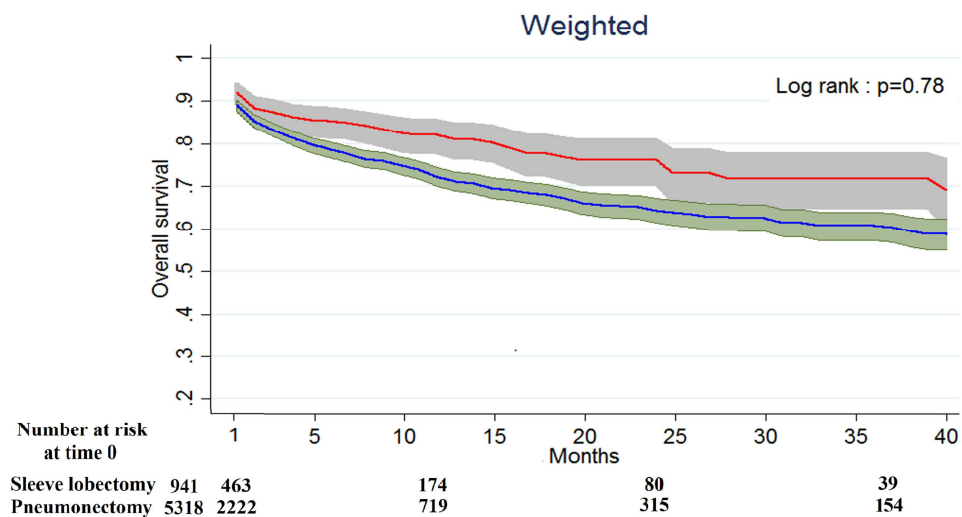
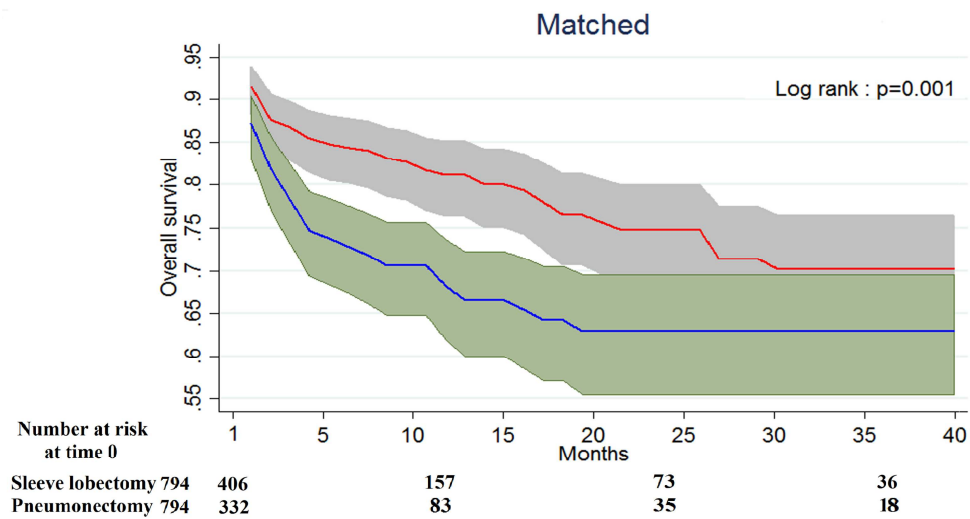
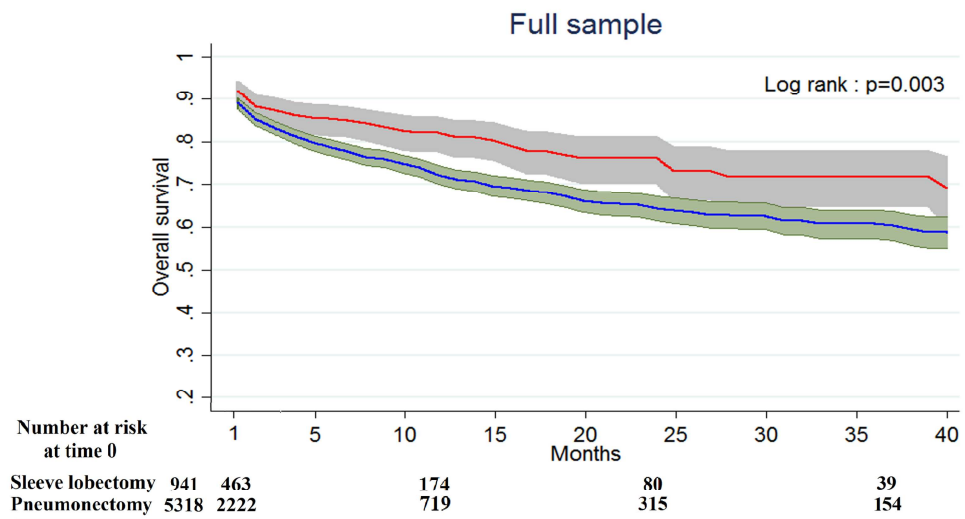
Figure 2: Kaplan-Meier plots of disease-free survival from the unmatched full sample (A), matched (B) and weighted (C) data for PN (continuous red line) and SL (discontinuous blue line).

Figure 3: Graph of propensity scores in the 2 groups of patients. Each bar represents the number of patients with the same propensity score in both groups.

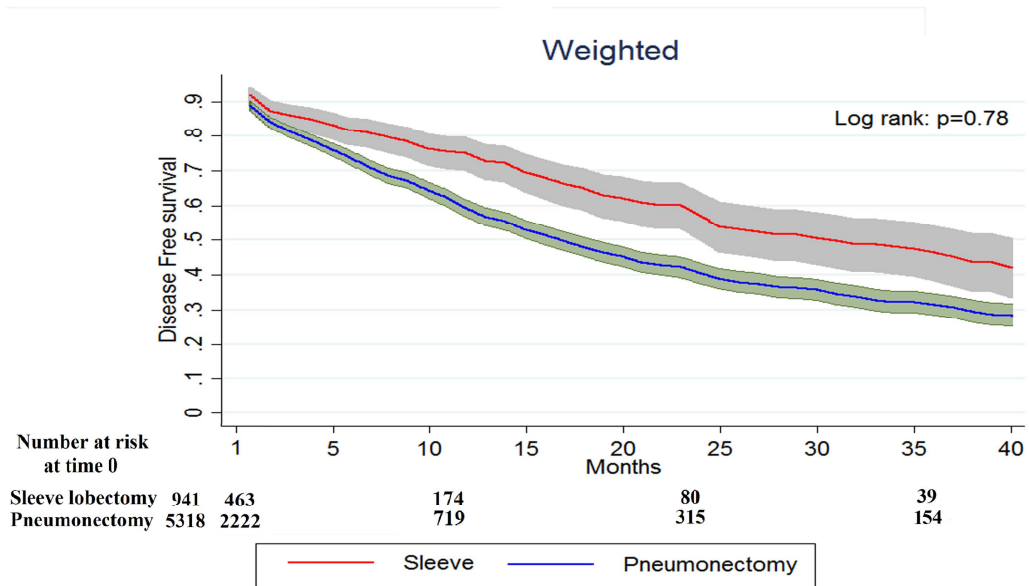
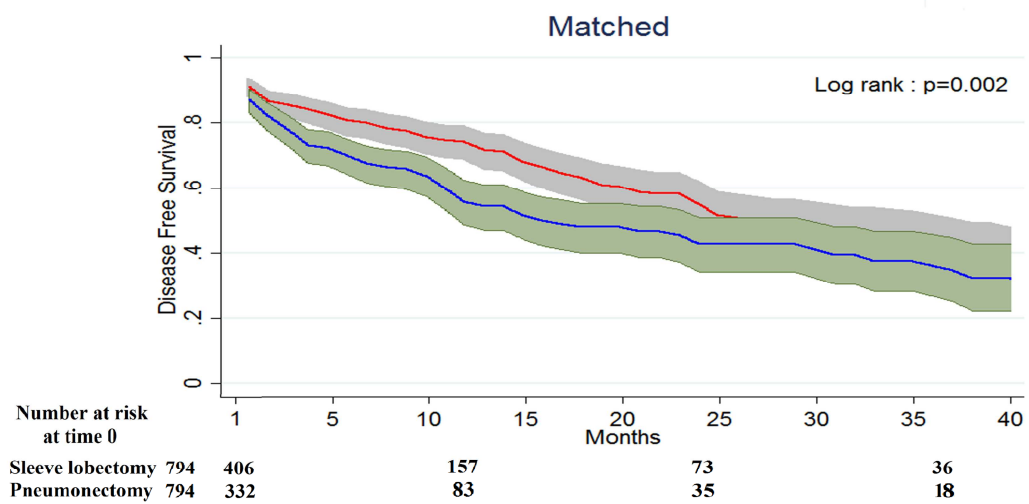
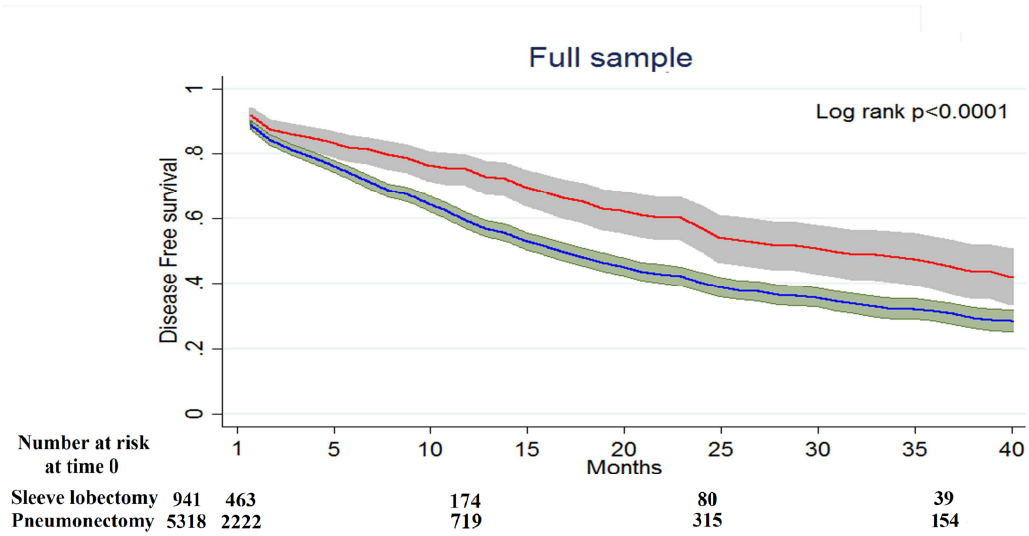
Video legend: this is a video of a double (vascular and bronchial) left upper lobectomy by thoracotomy.

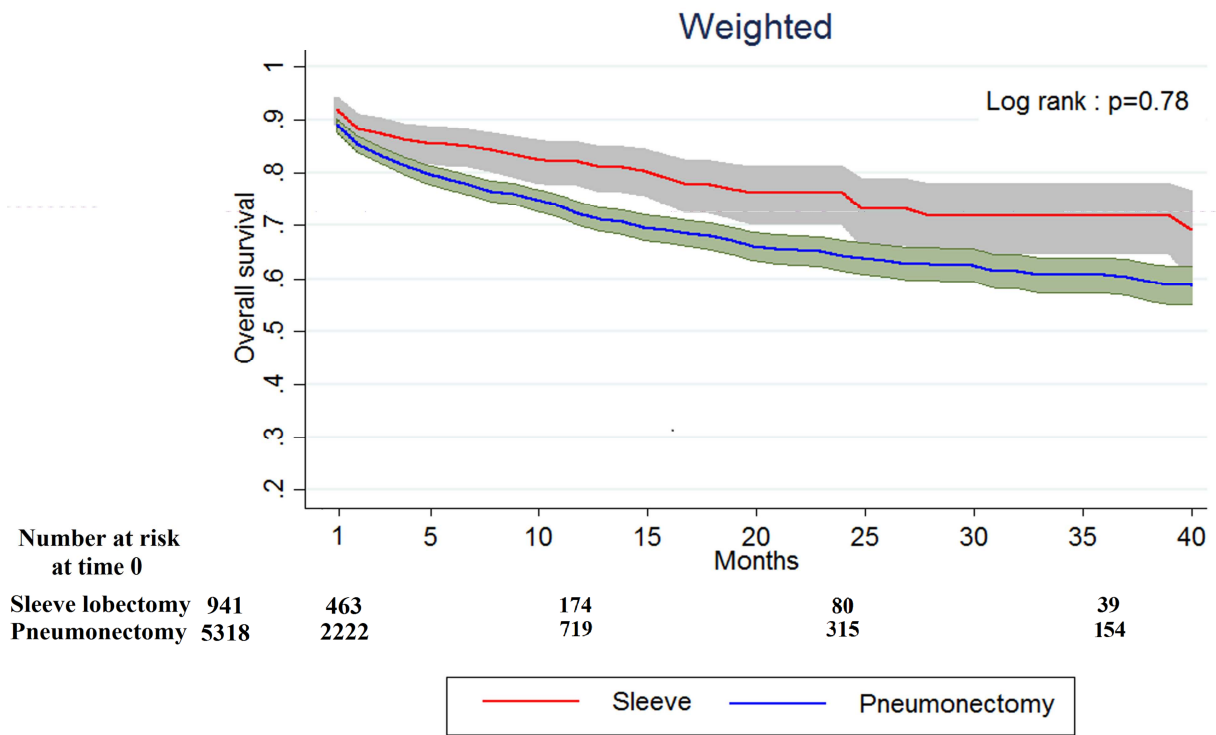


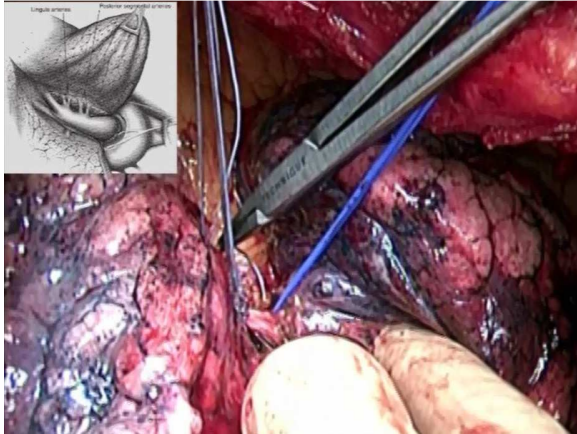
 Pneumonectomy  Sleeve



— Sleeve — Pneumonectomy







ACCEPTED MANUSCRIPT

Appendix 1

The authors would like to thank all the French thoracic surgeons who participated in Epithor and subsequently in this study in order to improve thoracic surgery quality: Dr Michel Alauzen (Montpellier), Dr Jean-François Andro (Quimper), Dr Maxime Aubert (Grenoble), Dr Jean Philippe Avaro (Marseille), Dr Patrick Bagan (Argenteuil), Dr François Bellenot (Cergy Pontoise), Vincent Blin (Vannes), Dr Philippe Boitet (Harfleur), Dr Laurent Bordigoni (Toulon), Professor Jacques Borrelly (Nancy), Professor Pierre-Yves Brichon (Grenoble), Dr Gilles Cardot (Boulogne sur Mer), Dr Jean Michel Carrie (Saint Jean), Dr François Clement (Besançon), Professor Pierre Corbi (Poitiers), Dr Michel Debaert (Lille), Dr Bertrand Debrueres (Ploemeur), Dr Jean Dubrez (Bayonne), Dr Xavier Ducrocq (Strasbourg), Dr Antoine Dujon (Bois Guillaume), Professor Pascal Dumont (Tours), Dr Philippe Fernoux (Chalon sur Saône), Professor Marc Filaire (Clermont-Ferrand), Dr Eric Frassinetti (Chambéry), Dr Gil Frey (Saint Etienne), Dr Dominique Gossot (Paris), Professor Gilles Grosdidier (Nancy), Dr Benoit Guibert (Lyon), Dr Olivier Hagry (Chalon sur Saône), Dr Sophie Jaillard (Lille), Dr Jean-Marc Jarry (Aix en Provence), Dr David Kaczmarek (Saint Etienne), Dr Yves Laborde (Pau), Dr Bernard Lenot (Saint Brieuc), Dr Francis Levy (Bordeaux), Dr Laurent Lombart (Béziers), Dr Eric Marcade (Saint Grégoire), Dr Jean Paul Marcade (La Rochelle), Dr Jean Marzelle (Créteil), Professor Gilbert Massard (Strasbourg), Dr Florence Mazeres (Bayonne), Dr Eric Mensier (Lille), Dr David Metois (Orléans), Dr J. L. Michaud/E Paris (Nantes), Dr Philippe Mondine (Brest), Dr Michel Monteau (Reims), Dr Jean-Michel Moreau (Nantes), Professor Jérôme Mouroux (Nice), Dr Antoine Mugniot (Nantes), Dr Pierre Mulsant (Lyon), Dr Nidal Naffaa (Avignon), Dr Gérard Pavy (Arras), Professor Christophe Peillon (Rouen), Professor François Pons (Percy), Professor Henri Porte (Lille), Professor Jean-François Regnard (Paris), Professor Marc Riquet (Paris), Dr Babak Sadeghi Looyeh (Morlaix), Professor Olivier Tiffet (Saint Etienne), Dr Bruno Tremblay (Meaux), Dr Jean Valla (Charenton le Pont), Professor Jean-François Velly (Pessac), Dr Bernard Wack (Metz), Dr Jean-Didier Wagner (Colmar), and Dr Didier Woelffe (Valenciennes).

Supplementary table 4

Sensitivity analysis to Hidden Bias for matching one-to-one Sensitivity (Mantel-Haenszel test statistic).

	$\hat{\Gamma}=1$ (assumption no hidden bias)		Conclusion
	Q _{mh}	P value	
Postoperative mortality	0.93	0.17	Caution, possible hidden bias
Postoperative complications			
Atelectasis	8.4	0.0001	No hidden bias
Pneumonia	4.4	0.0001	No hidden bias
Postoperative pulmonary complications	6.3	0.0001	No hidden bias
Arrhythmias	2.5	0.006	No hidden bias
Bronchopleural fistula	3.3	0.004	No hidden bias
Empyema	2.7	0.003	No hidden bias
Hemorrhage	4	0.0001	No hidden bias
Ventilation > 2 days	1.02	0.15	Caution, possible hidden bias
ARDS	0.9	0.16	Caution, possible hidden bias

ARDS: Acute Respiratory Distress Syndrome

Supplementary Table 5

Assessment of sensitivity to positivity violations for weighting method (excluding observations in regions of non-overlap).

	Weighting (n=4829)	
		P value
Postoperative mortality	0.74 (0.4-1.44)	0.4
Postoperative complications		
Atelectasis	0.13 (0.07-0.22)	0.0001
Pneumonia	0.55 (0.35-0.88)	0.01
Postoperative pulmonary complications	0.47 (0.33-0.67)	0.0001
Arrhythmias	1.86 (1.06-3.23)	0.03
Bronchopleural fistulae	2.4 (1.23-4.8)	0.01
Empyema	13 (3.5-48)	0.0001
Hemorrhage	1.8 (0.57-6)	0.3
Ventilation > 2 days	1.12 (0.6-2.1)	0.7
ARDS	1.05 (0.5-2.2)	0.9
Overall survival	0.99 (0.89-1.11)	0.9
Disease-Free survival	1.1 (1.01-1.2)	0.03

ARDS: Acute Respiratory Distress Syndrome