This is a pre-copyedited, author-produced version of an article accepted for publication in European Journal of Cardiothoracic Surgery following peer review. The version of recordJose Luis Recuero-Díaz, Iñigo Royo-Crespo, David Gómez de-Antonio, Sergi Call, Borja Aguinagalde, María Teresa Gómez-Hernández, Jorge Hernández-Ferrández, David Sánchez-Lorente, Julio Sesma-Romero, Eduardo Rivo, Nicolás Moreno-Mata, Raul Embun, on behalf of the Spanish Group of Video-assisted Thoracic Surgery (GEVATS), T*reatment and intention-to-treat propensity score analysis to evaluate the impact of video-assisted thoracic surgery on 90-day mortality after anatomical resection for lung cancer*, **European Journal of Cardio-Thoracic Surgery**, Volume 62, Issue 3, September 2022, ezac122 is available online at: https://doi.org/10.1093/ejcts/ezac122.

Treatment and intention-to-treat propensity score analysis to evaluate the impact of video-assisted thoracic surgery on 90-day mortality after anatomical resection for lung cancer

Authors

Jose Luis Recuero-Díaz[1†], Iñigo Royo-Crespo[1], David Gómez de-Antonio[2], Sergi Call[3], Borja Aguinagalde[4], María Teresa Gómez-Hernández[5], Jorge Hernández-Ferrández[6], David Sánchez-Lorente[7], Julio Sesma-Romero[8], Eduardo Rivo[9], Nicolás Moreno-Mata[10], and Raul Embun[1,11*†]

on behalf of the Spanish Group of Video-assisted Thoracic Surgery (GEVATS)

†These authors contributed equally to this work.

Affiliations

1. Department of Thoracic Surgery, Hospital Universitario Miguel Servet, Hospital Clínico Universitario Lozano Blesa, IIS Aragón, Zaragoza, Spain
2. Department of Thoracic Surgery, Hospital Universitario Puerta de Hierro Majadahonda, Madrid, Spain
3. Department of Thoracic Surgery, Hospital Universitari Mútua Terrasa, Universitat de Barcelona, Terrasa, Barcelona, Spain
4. Department of Thoracic Surgery, Hospital Universitario de Donostia, San Sebastián-Donostia, Spain
5. Department of Thoracic Surgery, Hospital Universitario de Salamanca, IBSAL, Universidad de Salamanca, Salamanca, Spain
6. Department of Thoracic Surgery, Hospital Universitario Sagrat Cor, Barcelona, Spain
7. Department of Thoracic Surgery, Hospital Clínic de Barcelona, Instituto Respiratorio, University of Barcelona, Barcelona, Spain
8. Department of Thoracic Surgery, Hospital General Universitario Alicante, Alicante, Spain
9. Department of Thoracic Surgery, Hospital Universitario Santiago de Compostela, Santiago de Compostela, Spain
10. Department of Thoracic Surgery, Hospital Ramón y Cajal, Madrid, Spain
11. Department of Surgery, Faculty of Medicine, University of Zaragoza, Zaragoza, Spain

***Corresponding author**

Raul Embun
Department of Thoracic Surgery
Hospital Universitario Miguel Servet
Paseo Isabel la Católica 1, Zaragoza 50009, Spain
Tel: +34-635492179
Email: raulembun@gmail.com

Abstract

OBJECTIVES: The aim of this study was to know the treatment effect of video-assisted thoracic surgery (VATS) on 90-day mortality after anatomical lung resection based on a nationwide cohort.

METHODS: This is a multicentre prospective cohort of 2721 anatomical resections for lung cancer from December 2016 to March 2018. Treatment and intention-to-treat (ITT) analyses were performed after inverse probability score weighting and different propensity score matching algorithms. Covariate balance was assessed by standardized mean differences. The estimators reported were the average treatment effect, the average treatment effect on the treated and odds ratios after conditional logistic models with 95% confidence intervals. The unconfoundedness assumption was evaluated by sensitivity analysis for average treatment effect (c-dependence) and average treatment effect on the treated (Γ).

RESULTS: VATS was the initial approach in 1911 patients (70.2%), though 273 cases (14.3%) had to be converted to thoracotomy. Ninety-day mortality rates were: treatment analysis (VATS 1.16% vs open 3.9%, P < 0.001), ITT analysis (VATS 1.78% vs open 3.36%, P = 0.012). After inverse probability score weighting and propensity score matching, in the treatment analysis, VATS meant absolute risk reductions between 2.25% and 2.96% and relative risk reductions between 65% and 70% (OR = 0.34, 95% confidence interval 0.15–0.79, all P-values <0.004). However, all the estimators turned out to be non-significant in the ITT analyses. A high sensitivity to unobservable confounders was proved (c-dependence 0.135, Γ = 1.5).

CONCLUSIONS: VATS can reduce the risk of 90-day mortality after anatomical lung resection. However, the implications of conversion to thoracotomy, comparing ITT versus treatment analysis, and the potential impact of hidden bias should deserve further attention in the future.

Keywords: Video-assisted thoracic surgery, 90-Day mortality, Intention-to-treat analysis, Thoracic surgery, Anatomical lung resection, Lung cancer

---

Abbreviations

ATE - Average treatment effect
ATT - Average treatment effect on the treated
CIs - Confidence intervals
GEVATS - Spanish Group of Video-assisted Thoracic Surgery
ITT - Intention-to-treat
IPSW - Inverse probability score weighting
PSM - Propensity score matching

SMD - Standardized mean differences
VATS - Video-assisted thoracic surgery

Introduction

Although there are multiple retrospective series that have shown a lower rate of complications and postoperative stay in patients operated on by video-assisted thoracic surgery (VATS), there are more discrepancies on the impact that the surgical approach could have on postoperative mortality. In this regard, a recent randomized clinical trial, designed to evaluate short-term and oncologic efficacy of VATS, failed to demonstrate differences in postoperative mortality between VATS and thoracotomy [1].

Regarding postoperative mortality, recent series have shown that 90-day mortality could double 30-day or in-hospital mortality after lung resection [2–4]. Although most series comparing VATS with thoracotomy do not mention 90-day mortality, some important studies have not shown significant differences in either in-hospital, 30-day or 90-day mortality [5–7]. A recent publication from the ESTS database did show significant differences in mortality at discharge (VATS 1% vs thoracotomy 1.9%, P = 0.020), not mentioning mortality at 90 days [8].

Few studies have performed an intention-to-treat (ITT) analysis comparing VATS and thoracotomy for lung resection, and even less have compared the results obtained depending on the strategy of analysis performed [9, 10]. However, since a non-negligible proportion of patients undergoing VATS must be finally converted to thoracotomy, ITT analysis seems to be the most appropriate strategy to evaluate treatment effects related to VATS in a real scenario [11].

In view of the conflicting results and the underreported ITT analysis in the literature, the objective of our study was to determine the impact of the surgical approach on 90-day mortality, comparing treatment and ITT analysis in patients who underwent and anatomical lung resection for lung cancer in the national cohort by the Spanish Group of Video-assisted Thoracic Surgery (GEVATS) [12].

Materials and Methods

Ethical statement

This project was approved by all the local ethics committees and informed consent was obtained from the recruited patients to use their clinical data for scientific purposes (Approval by Ethics Committee of Aragon Health Research Institute on 20 May 2015 PI15/0072).

Data source

The GEVATS of the Spanish Society of Thoracic Surgery is a prospective voluntary multicentre observational study with a total of 33 Thoracic Surgery Department participating. The Centres were not selected based on their experience in VATS or any other criterion. The cohort included patients who underwent an anatomical lung resection from 20 December 2016 to 20 March 2018. The GEVATS objectives were to know the impact of surgical approach on short- and long-term outcomes. The method of the GEVATS, including sample size justification and the audit process performed, was recently published [12].

Patient allocation into VATS or thoracotomy depended on clinical practice, experience and beliefs of each of the participating surgeons and departments.

In this prospective observational cohort study, we tried to specifically elucidate the impact of the surgical approach on 90-day mortality after anatomical lung resection for lung cancer. For this purpose, those patients with a diagnosis other than lung carcinoma and those who underwent pneumonectomy or extended lung resection were excluded. Extended resection was considered in case of chest wall, diaphragm or sleeve resection. Our manuscript is reported according to the STROBE recommendations and ESTS Statistical Primer for propensity score analysis.

Statistical analysis

A double type of analysis was carried out, treatment and ITT, depending on the need of conversion from VATS to open throughout the procedure. The association of the surgical approach (open versus VATS) as the treatment variable, with baseline oncologic and surgical variables that could influence on the outcome variable (90-day mortality) and the approach to be chosen, was analysed by two-tailed statistical hypothesis testing, using Mann–Whitney and Chi-square tests and standardized mean differences (SMD). Those variables with a P-value of less than 0.2 and/or standardized differences greater than 0.1 were the covariates used to build the propensity score to correct for selection bias.

Missing data were dealt by casewise deletion analysis when less than 5% of patients had incomplete registries.

The propensity score was estimated by a logit model and the overlap assumption was assessed on density plots for treatment and ITT analysis. The treatment effects were evaluated based on the inverse probability score weighting (IPSW) and the propensity score matching (PSM) through the nearest-neighbor method with and without replacement, using different calliper widths (0.035, 0.05 and 0.1) and matching ratios (1:1, 1:2 and 1:3).

Covariate balanced was assessed by SMD, before and after weighting or matching. SMD less than 0.1 or 0.05 were considered good or excellent, respectively, to exclude residual imbalance [14]. Balance of covariates was displayed on dot plots for IPSW and PSM, separately.

The treatment effects were estimated by weighted mean and matching outcome models reporting the average treatment effect (ATE), based on the difference in potential outcome means, and the average treatment effect on the treated (ATT). ATE and ATT were reported as absolute risk reductions (IPSW and PSM) and relative risk reductions (IPSW). In addition, in case of PSM 1:1 without replacement and calliper 0.035, a conditional logistic fixed-effects regression model was conducted. In this case, treatment effects were reported as odds ratios. The impact of surgeon experience in VATS procedures (≤50 versus >50 cases), surgeon seniority (resident versus faculty <10 years versus faculty 10–20 years versus faculty >20 years), surgical volume and VATS rate by department (discrete variables) were used to adjust the odds ratios reported by the conditional logistic models previously described. To compute surgical volume and VATS rate by institution throughout the 15-month recruitment period, we considered the reports submitted by the Heads of the Administrative Departments from each Institution. These reports were used in the audit process we previously published [12]. 95% confidence intervals (CIs) were calculated from robust standard errors and P-values less than 0.05 were considered statistically significant.

The conditional partial dependence method proposed by Masten et al. was used to evaluate the sensitivity of conclusions about the ATE. Bounds on the ATE given a set of c-dependence values (between 0 and 1) and the breakdown point (maximum value of the c-dependence parameter under which the conclusion still holds) were reported. In addition, the impact of hidden bias on the ATT estimator after PSM (1:1, calliper 0.035, no replacement) was assessed with the bounding approach proposed by Rosenbaum. The Γ parameter and corresponding P-values were used to measure the sensitivity of ATT to unobservable confounders.

The Treatment Effects Suite in Stata/MP 16.0 and the Stata packages Stddiff, Psmatch2, Calipmatch, Tesensitivity and Mhbounds were used for the statistical analysis. Tablau Desktop 2020.3.1 was used for plot representation of the covariate balance.

Results

A total of 3533 patients were recruited, including 1917 VATS cases (54.3%). After exclusion of patients with a diagnosis different to lung cancer (448 patients, 12.7%), pneumonectomy (236 patients, 6.7%) and extended resection (165, 4.7%), 2721 patients (77% of the entire GEVATS cohort) met the inclusion criteria. The types of resections included were: 2444 lobectomies (90%), 111 bilobectomies (4%) and 166 anatomical segmentectomies (6%).

VATS was the initial approach in 1911 patients (70.2%), representing the treatment group in the ITT analysis. However, 273 cases (14.3%) had to be converted to thoracotomy and, therefore, the VATS arm in the treatment analysis consisted of 1638 patients (60.2%). The unadjusted analysis showed an important association between in-hospital mortality and 90-day mortality with the surgical approach performed (Table 1). The outcome variable (90-day mortality) was missing in 11 cases (0.4%) that were not considered in the analysis of treatment effects. The percentages of missing values were negligible for all the confounders except for DLCO. The main analysis excluded DLCO as confounder, so 14 variables were used to build the propensity score in a complete-case analysis since only 1.8% of the patients had missing values in some of the covariates included (Table 2). The propensity score distribution to prove the overlap assumption is shown in a density plot (Figure 1). Only 9 patients in the VATS group (0.6%) had a propensity score higher than the maximum propensity score value in the open group, while only 2 patients in the open group (0.2%) had a propensity score lower than the minimum value in the VATS group.

After IPSW, the ratios of VATS to thoracotomy were 1339:1324 patients in the treatment analysis and 1347:1317 patients in the ITT analysis. Covariate balance was proved (Figure 2).

PSM 1:1 without replacement and a calliper width of 0.035 yielded 872 matched pairs in the treatment analysis and 705 matched pairs in the ITT analysis. Most of the covariates reached an excellent balance after matching (Table 3). The proportions of cases matched to 1 control were 54.2% (treatment analysis) and 37.6% (ITT analysis).

After PSM with replacement, 2664 patients (1608 VATS and 1056 open) in the treatment analysis and 2663 patients (1877 VATS and 786 open) in the ITT analysis were matched to 1, 2 or 3 counterparts. Independently of the matching ratios and calliper widths, most of the covariates reached an excellent balance (Figure 3).

Based on the treatment analysis, VATS was consistently associated to a lower 90-day mortality rate according to the ATE and ATT estimators, whatever the propensity score technique or

algorithm used. However, in the case of ITT analysis, the still lower mortality after VATS was associated to non-significant ATE and ATT, in terms of absolute and relative risk reductions, in all the cases (Table 4).

In the conditional logistic model after PSM 1:1 (no replacement and calliper width 0.035), the results obtained after treatment analysis (OR = 0.38; 95% CI 0.20–0.73; P = 0.004) and ITT analysis (OR = 0.71; 95% CI 0.38–1.33; P = 0.283) were consistent with the ATE and ATT estimators. After adjusting the fixed-effect models by surgeon experience (number of VATS procedures and seniority) and department experience (surgical volume and VATS rate) the odds ratios were similar to those in the non-adjusted matched treatment analysis (OR = 0.34; 95% CI 0.15–0.79; P = 0.012) and lower, though still non-significant, in the ITT analysis (OR = 0.41; 95% CI 0.13–1.25; P = 0.117).

In the sensitivity analysis, the close to zero c-dependence values under which the conclusions still held (treatment analysis 0.135 and ITT analysis 0.09) meant a high sensitivity of the ATE estimator to hidden bias (Figure 4). In addition, the maximum Γ value at which significant P-values were obtained (Γ = 1.5) showed that the ATT estimation of this study is sensitive to bias (unobserved variables) able to increase the odds of receiving VATS in more than 50%.

In the main analysis, we did not consider DLCO because of the high rate of missing values (15.2%). However, in a secondary analysis including DLCO as a covariate to build the propensity score, the ATE and ATT estimators after IPSW were equivalent to those obtained in the main analysis. Moreover, the sensitivity analysis showed an even higher potential impact of unmeasured or unobservable confounders (Γ = 1.15).

Discussion

Our main finding was that VATS can reduce 90-day mortality after an anatomical lung resection for lung cancer. However, this beneficial effect decreased and turned out to be non-significant when an ITT analysis was performed. Noteworthy, these findings were consistent along all the propensity score algorithms carried out. Finally, the high sensitivity of our analysis to hidden bias highlights that the potential impact of unmeasured and unobservable confounders in observational studies comparing VATS and open surgery could be determinant.

The results of the few studies comparing the impact of the surgical approach on 90-day mortality lead to contradictory conclusions. Limited statistical power secondary to small sample sizes and surprisingly low mortality rates could partly explain non-conclusive findings [1, 6, 7]. However, when analysing the results obtained by large institutional cohorts, significant statistical differences could only correspond to clinically trivial treatment effect [15].

In the thoracic surgery literature, useful recommendations have been made to improve reporting based on the propensity score analysis, since this analytic approach is not as straightforward as regression analysis [16]. Considering the recommendations by the European Society of Thoracic Surgeons in its statistical primer report, we carried out the 2 advised methods, IPSW and PSM with calliper [13]. In addition, to assess the consistency of our results, we put into practice different matching algorithms (treated to control ratios, calliper widths and replacement versus no replacement matching).

The use of 2 propensity scores techniques comparing outcomes after VATS or open lung resection is an exception [10]. Pagès et al. published the results by a French nationwide study

evaluating the impact of surgical approach on short- and long-term outcomes after lobectomy for lung cancer. In this large cohort (n = 24,811) recruited throughout an 8-year period, only 4.9% of the cases were operated on by VATS. It is noteworthy that the significance level obtained in a few outcomes differed between the propensity score strategies used (IPSW versus PSM). However, regarding 30-day mortality, both IPSW and PSM failed to demonstrate a significant reduction after VATS (PSM OR = 0.89, 95% CI 0.45–1.81; IPSW OR = 0.74, 95% CI 0.37–1.45).

The importance of an ITT strategy when evaluating the impact of the surgical approach was highlighted in one of the first meta-analyses published comparing VATS and thoracotomy [17]. However, the information about conversion from VATS to open surgery is not always registered, even in case of national and international registries [6, 8, 10, 18, 19]. Consequently, ITT analysis is not a wide practice in the thoracic surgery literature and, therefore, the benefit conferred to VATS when evaluating treatment effects could be misleading. In fact, the higher the rate of conversion from VATS to open the higher the potential discrepancy between treatment and ITT analyses. In this regard, despite a higher rate of VATS in our cohort (60.2%) than the one reported by many nationwide and multicentre registries, the proportion of conversion we are reporting (14.3%) is notably superior compared to some recognized series ranging from 2% to 9% [11, 20, 21]. Risk of conversion in a recent metanalysis was 9.6% (95% CI: 6.6–13.9%) [22]. We cannot conclude about the reasons for our higher conversion rate. However, nationwide representativeness, data audit that involved all the participating centres, or just a matter of disposition of surgeons, more prone to starting the procedure by VATS in our country, could be some of the hypotheses. In addition, although anyone could expect that higher VATS rates are parallel to lower conversion rates because of experience acquisition, this reasoning could be conflicting depending on the scenario and the type of cohort in question, for example single centre versus multicentre. In this regard, our relatively high conversion rate could simply represent a consequence of our directly proportional high VATS rate.

Although some studies argue that conversion from VATS to open does not entail a surgical failure, other series have associated conversion to worse outcomes in terms of postoperative morbidity [20, 23]. Our study would support a detrimental effect associated to conversion. However, more detailed and specific analyses comparing converted VATS to straight thoracotomy could generate insightful knowledge about the implications of conversion itself and the most appropriate disposition towards starting the procedure by VATS.

ITT analysis is claimed as the gold standard in randomized control trials to reflect a pragmatic clinical scenario, to maintain prognostic balance generated from the original randomization and to give an unbiased estimate of treatment effect. However, we considered important at this stage when most of the studies dealing with VATS are still reporting their results based on a treatment analysis, to include this less conservative strategy of analysis so that we could still compare our findings with the coetaneous literature. To our knowledge, this is the first manuscript that evaluate the impact of the surgical approach on short-term mortality after lung resection, comparing a treatment and ITT analysis.

Limitations

The main drawback in multicentre voluntary registries is related to selection bias and data quality. Although the details of our audit were previously reported, we cannot reject residual bias at this stage [12].

As previously published, to calculate our sample size, we considered 2% absolute risk difference in 90-day mortality as relevant in practice (4% vs 2%), and an expected 25% of cases by VATS [12]. Because of the underestimated 90-day mortality rate difference between VATS and Open (1.16% vs 3.9%) and the higher-than-expected VATS rate (60%), the statistical power of our 'positive' conclusion in the treatment analysis (99%) is higher than the conventional 80%. However, in the ITT analysis, the lower difference in mortality between both approaches (1.78% vs 3.36%) and the lower proportion of control patients (29%) could be the reasons for a limited power (69%) of our 'negative' results.

Missing values are another cornerstone in observational studies. Although DLCO was the only covariate with a missing rate of higher than 2% (415/2721 patients 15%), this respiratory parameter has been proved as one of the most determinant risk factors after lung resection, even in our own cohort [24–26]. Despite equivalent results in a post hoc analysis based on the subgroup of patients with this value present, confounding bias cannot be excluded.

Despite the national representativeness of the GEVATS cohort, since it would have included 50% of the anatomical lung resections performed in Spain over the 15-month recruitment period, the high audited rate of VATS registered (60%) could compromise the exportation of our results to other national or local scenarios with an important discrepancy in VATS implementation [12]. However, the equipoise in our study in terms of the proportion of patients in each group compared to other series makes it less likely that either our open group represents a strongly selected collection of challenging cases or our VATS group a strongly selected collection of favourable cases [6, 10, 27].

PSM is considered to jeopardize the generalization of its conclusions when matched sample size is lower than 50% of the original sample. In our study by carrying out PSM 1:1 without replacement and calliper width 0.035, the proportion of matched sample was 64% and 52% after treatment and ITT analysis, respectively. Nevertheless, the results obtained after such PSM algorithms were equivalent to those reached after making use of all the sample with other propensity score techniques, which seems to preserve our conclusion at a more general setting.

Finally, hidden bias is a ubiquitous problem in every observational study, even in case of a rigorous propensity score analysis, and sensitivity test tries to get a better understanding of this obstacle to obtain reliable estimators. However, this type of analysis is exceptional in thoracic surgery [10]. The study published by Pagès et al. from the Epithor database reported a high sensitivity to unobservable confounders when estimating postoperative death ($\Gamma = 1.6$). Our results were very similar ($\Gamma = 1.5$), even after including DLCO as a covariate in the propensity score analysis ($\Gamma = 1.2$, the closer to 1 the more sensitive our estimations to hidden bias). Consequently, the reliability of our estimations should be cautiously admitted until more robust evidence.

Conclusion

Having in mind the exclusion criteria previously referred (pneumonectomy and extended resection), VATS was consistently shown to reduce 90-day mortality by 65–70% after an anatomical lung resection for lung cancer in the GEVATS cohort. However, drawn on the ITT analysis including patients converted to open surgery into the VATS group, this benefit decreased to the extent of obtaining statistically non-significant, though potentially

underpowered, differences. Until upcoming evidence comparing converted VATS to straight thoracotomy, an effort should be made to optimize the beneficial effect of VATS on short-term mortality after an anatomical resection for lung cancer. In the analysis of hidden bias, we demonstrated that the treatment effect estimators were highly sensitive to unobservable confounders, which could only be overcome in a sufficiently large multicentre randomized trial.

Acknowledgements

We thank all the local researchers who contributed with patient recruitment and data entry.

Coordinator: Raul Embun (University Hospital Miguel Servet).

Scientific committee: David Gómez de Antonio (University Hospital Puerta de Hierro Majadahonda, Madrid); Sergi Call (MútuaTerrasa University Hospital, University of Barcelona, Terrasa, Barcelona); Nicolás Moreno-Mata (Ramón y Cajal University Hospital, Madrid); Marcelo F. Jiménez (Salamanca University Hospital, University of Salamanca, IBSAL, Salamanca); Miguel Congregado (Virgen Macarena University Hospital, Seville); and Sergio Bolufer-Nadal (General University Hospital of Alicante).

Local researchers: Jose Luis Recuero and Iñigo Royo Crespo (University Hospital Miguel Servet and Lozano Blesa); Borja Aguinagalde and Iker López Sanz (Donostia University Hospital, San Sebastián-Donostia); Sergio Amor-Alonso and Francisco Javier Moradiellos-Díez (Quironsalud Madrid University Hospital, Madrid); Miguel Jesús Arrarás (Valencian Oncology Institute Foundation, Valencia); Ana Isabel Blanco Orozco (University Hospital Virgen del Rocío, Sevilla); Marc Boada and David Sánchez (Hospital Clinic of Barcelona, Respiratory Institute, University of Barcelona, Barcelona); Alberto Cabañero Sánchez (Ramón y Cajal University Hospital, Madrid); Isabel Cal Vázquez and Ramón Moreno-Balsalobre (University Hospital La Princesa, Madrid); Ángel Cilleruelo Ramos (University Clinical Hospital, Valladolid); Silvana Crowley Carrasco (University Hospital Puerta de Hierro Majadahonda, Madrid); Elena Fernández-Martín and Florentino Hernando Trancho (Hospital Clínico San Carlos, Madrid); Santiago García-Barajas and Cipriano López García (University Hospital of Badajoz, Badajoz); Maria Dolores García-Jiménez (University Hospital of Albacete, Albacete); Jose María García-Prim and Eduardo Rivo (Santiago de Compostela University Hospital, Santiago de Compostela); Jose Alberto Garcia-Salcedo (12 de Octubre University Hospital, Madrid); Juan José Gelbenzu-Zazpe and María Elena Ramírez-Gil (Complejo Hospitalario of Navarre, Pamplona); Carlos Fernando Giraldo-Ospina and Roberto Mongil Poce (Hospital Regional Universitario, Málaga); María Teresa Gómez Hernández (Salamanca University Hospital, University of Salamanca, IBSAL, Salamanca); Jorge Henández and Juan José Fibla Alfara (Sagrat Cor University Hospital, Barcelona); Jennifer D. Illana Wolf (Puerta del Mar Hospital, Cádiz); Alberto Jauregui Abularach (Vall d'Hebron University Hospital, Barcelona); Unai Jiménez and Rafael Rojo-Marcos (University Hospital Cruces, Bilbao); Néstor J. Martínez-Hernández (La Ribera University Hospital, Alcira, Valencia); Elisabeth Martínez-Téllez and Juan Carlos Trujillo Reyes (Santa Creu and Sant Pau Hospital, Autonomous University of Barcelona, Barcelona); Lucía Milla Collado (Hospital Arnau de Vilanova, Lleida); Sergio B. Moreno Merino (Virgen Macarena University Hospital, Seville); Carme Obiols (MútuaTerrasa University Hospital, University of Barcelona, Terrasa, Barcelona); Florencio Quero-Valenzuela (Virgen de las Nieves Hospital, Granada); Ricard Ramos-Izquierdo (University Hospital of Bellvitge, Hospitalet de Llobregat, Barcelona); Alberto Rodríguez-Fuster (Hospital del Mar, Hospital del Mar Medical Research Institute, Barcelona); Laura Sanchez Moreno (University

Hospital Marqués de Valdecilla, Santader); Carlos Simón (Gregorio Marañón University Hospital, Madrid); and Julio Sesma Romero (General University Hospital of Alicante).

Funding

GEVATS is supported by Ethicon and Spanish Society of Thoracic Surgery.

Conflict of interest

The authors declare no conflicts of interest.

Data Availability Statement

The data underlying this article will be shared on reasonable request to the corresponding author.

Author contributions

Jose Luis Recuero-Díaz: Conceptualization; Methodology; Resources; Writing—original draft.
Iñigo Royo-Crespo: Resources; Visualization; Writing—review & editing.
David Gómez de-Antonio: Resources; Writing—review & editing.
Sergi Call: Resources; Writing—review & editing.
Borja Aguinagalde: Resources; Writing—review & editing.
María Teresa Gómez-Hernández: Resources; Writing—review & editing.
Jorge Hernández-Ferrández: Resources; Writing—review & editing.
David Sánchez-Lorente: Resources; Writing—review & editing.
Julio Sesma-Romero: Resources; Writing—review & editing.
Eduardo Rivo: Resources; Writing—review & editing.
Nicolás Moreno-Mata: Resources; Writing—review & editing.
Raul Embun: Conceptualization; Formal analysis; Methodology; Project administration; Resources; Supervision; Visualization; Writing—original draft.

Tables and Figures

Table 1: Unadjusted analysis of in-hospital mortality and 90-day mortality related to the surgical approach performed, drawn on treatment and intention-to-treat analysis

	Treatment analysis			**Intention-to-treat analysis**		
	VATS, n/N (%)	Open, n/N (%)	P-Value	VATS, n/N (%)	Open, n/N (%)	P-Value
In-hospital mortality	11/1638 (0.67)	18/1083 (1.66)	0.019	17/1911 (0.89)	12/810 (1.48)	0.169

	Treatment analysis			Intention-to-treat analysis		
90-Day Mortality	19/1633 (1.16)	42/1077 (3.9)	<0.001	34/1906 (1.78)	27/804 (3.36)	0.012

Note: Ninety-day mortality rate in case of VATS converted to thoracotomy was 5.49% (15/273 patients) compared to 3.36% in case of straight thoracotomy (27/804 patients).

VATS: video-assisted thoracic surgery.

Table 2: Univariate analysis showing the association between baseline, oncologic and surgical variables with the surgical approach

Variable	VATS, n = 1638	Open, n = 1083	Standardized differences	P-Value	Missing (%)
Baseline variables					
Male	1103 (67.3)	782 (72.3)	-0.107	0.006	0.04
Age	67 (60–73)	66 (59–73)	0.079	0.195	0
BMI	26.6 (23.7–29.6)	26.5 (23.8–29.7)	-0.013	0.79	2.02
Smoking history	1393 (86.2)	931 (87.6)	-0.042	0.284	1.51
High blood pressure	768 (47)	482 (44.5)	0.049	0.207	0.11
Ischaemic heart disease	153 (9.3)	104 (9.6)	-0.008	0.819	0
Arrhythmia	132 (8.1)	96 (8.9)	-0.029	0.453	0.04
Cerebrovascular accident	88 (5.4)	60 (5.6)	-0.007	0.846	0.04
Diabetes	301 (18.4)	214 (19.8)	-0.035	0.361	0.04
Creatinine >2 mg/dl	52 (3.2)	26 (2.4)	0.047	0.235	0.04
History of thoracic surgery	64 (3.9)	69 (6.4)	-0.111	0.004	0
MRC dyspnoea ≥1	564 (34.4)	439 (40.6)	-0.127	0.001	0.04
FEV1	90 (76–104)	85 (73–97)	0.259	<0.001	0.96
DLCO	82 (69–98)	80 (67–92)	0.147	<0.001	15.25
ASA 3–4	879 (53.7)	642 (59.5)	-0.115	0.003	0.26

Variable	VATS, n = 1638	Open, n = 1083	Standardized differences	P-Value	Missing (%)
Oncologic/surgical variables					
Tumour size	21 (15–32)	30 (16–45)	-0.434	<0.001	0.44
Central tumour	410 (25.1)	523 (48.3)	-0.497	<0.001	0.11
cN2–3 (CT scan)	107 (6.5)	144 (13.3)	-0.227	<0.001	0.15
cN2–3 (PET scan)	161 (9.9)	205 (19)	-0.261	<0.001	0.15
Neoadjuvancy	70 (4.3)	123 (11.4)	0.266	<0.001	0
Right haemithorax	1007 (61.5)	636 (58.7)	0.056	0.151	0
Bilobectomy	21 (1.3)	90 (8.3)	-0.333	<0.001	0
Segments	3 (3–5)	4 (3–5)	-0.112	0.003	0.04

Notes:

- History of thoracic surgery: thoracic surgery performed previously under general anaesthesia.
- Modified MRC dyspnoea scale.
- Central tumours were considered when located in the inner one-third of the hemithorax.
- Functioning segments resected were evaluated by CT scan.
- Bold: variables with a P-value of <0.2 and/or standardized differences >0.1 were included in a logit model to calculate the propensity scores.

BMI: body mass index; MRC: Medical Research Council; VATS: video-assisted thoracic surgery.

Table 3: Number of matched patients and covariate balance after propensity score matching 1:1 without replacement and calliper width 0.035

Variable	Treatment analysis matched 872 VATS:872 open			Intention-to-treat analysis matched 705 VATS:705 open		
	Open	VATS	SMD	Open	VATS	SMD
	n (%) / mean (SD)	n (%) / mean (SD)		n (%) / mean (SD)	n (%) / mean (SD)	

Variable	Treatment analysis matched 872 VATS:872 open			Intention-to-treat analysis matched 705 VATS:705 open		
Male	623 (71.4)	618 (70.9)	-0.012	501 (70.9)	502 (71)	-0.003
Age	65.5 (10.4)	66.1 (9.43)	0.055	65 (10.4)	65.4 (9.7)	0.043
History of thoracic surgery	53 (6.1)	48 (5.5)	-0.024	47 (6.6)	46 (6.5)	-0.006
MRC dyspnoea ≥1	347 (39.8)	351 (40.3)	-0.009	288 (40.7)	289 (40.9)	-0.003
FEV1	86.8 (18.6)	87.4 (20.5)	0.031	84.9 (19)	85.9 (19.5)	0.053
ASA 3–4	506 (58)	530 (60.8)	-0.056	413 (58.4)	438 (62)	-0.072
Size tumour	29.5 (19.6)	27.9 (17.5)	-0.079	32.3 (21.3)	30.6 (20.1)	-0.082
Central tumour	357 (40.9)	335 (38.4)	-0.051	341 (48.2)	333 (47.1)	-0.022
N2–3 (CT scan)	84 (9.6)	85 (9.7)	-0.003	91 (12.9)	90 (12.7)	-0.004
N2–3 (PET scan)	131 (15)	122 (14)	-0.029	129 (18.2)	125 (17.7)	-0.014
Neoadjuvancy	65 (7.5)	59 (6.8)	-0.026	76 (10.7)	66 (9.3)	-0.047
Right haemithorax	503 (57.7)	511 (58.6)	-0.018	417 (59)	431 (61)	-0.040
Bilobectomy	22 (2.5)	21 (2.4)	-0.007	42 (5.9)	38 (5.4)	-0.024
Functioning segments resected	3.65 (1.31)	3.67 (1.22)	0.014	3.67 (1.36)	3.7 (1.32)	0.017

Notes:

- History of thoracic surgery: thoracic surgery performed previously under general anaesthesia.

- Modified MRC dyspnoea scale.
- Central tumours were considered when located in the inner one-third of the hemithorax.

MRC: Medical Research Council; SD: Standard deviation; SMD: standardized mean differences; VATS: video-assisted thoracic surgery.

Table 4: Treatment effect of video-assisted thoracic surgery on 90-day mortality after inverse probability score weighting in the treatment and intention-to-treat analyses

	POmeans (95% CI)	**ATE**		**ATT**	
		Absolute risk reduction (95% CI)	Relative risk reduction (95% CI)	Absolute risk reduction (95% CI)	Relative risk reduction (95% CI)
Treatment analysis					
Open	3.65% (2.48–4.82)				
VATS	1.11% (0.6–1.61)	2.54% (1.27 to 3.81)	69.6% (52.7 to 86.5)	2.25% (0.88 to 3.62)	65.6% (45.7 to 85.5)
P-Values		<0.001	<0.001	0.001	<0.001
Intention-to-treat analysis					
Open	2.87% (1.70–4.05)				
VATS	2.16% (1.29–3.03)	0.72% (-0.74 to 2.17)	24.9% (-0.68 to 18.1)	0.83% (-0.53 to 2.18)	31.3% (-7.6 to 70.2)
P-Value		0.337	0.257	0.232	0.115

ATE: average treatment effect; ATT: average treatment effect on the treated; CI: confidence interval; POmeans: potential outcome means; VATS: video-assisted thoracic surgery.

Figure Legends

Figure 1: Distribution of the propensity scores in the (a) treatment and (b) intention-to-treat analyses.

Figure 2: Balance plot of covariates after inverse probability score weighting (treatment and intention-to-treat analysis).

Figure 3: Balance plot of covariates after propensity score matching (intention-to-treat analysis).

Figure 4: Average Treatment Effect bounds based on incremental c-dependence parameter values. (Vertical dot-dashed lines indicate the breakdown points.)

References

1. Long H, Tan Q, Luo Q, Wang Z, Jiang G, Situ D et al. Thoracoscopic surgery versus thoracotomy for lung cancer: short-term outcomes of a randomized trial. Ann Thorac Surg 2018;105:386–92.
2. Pezzi CM, Mallin K, Mendez S, Gay EG, Putnam JB. Ninety-day mortality after resection for lung cancer is nearly double 30-day mortality. J Thorac Cardiovasc Surg 2014;148:2269–78.
3. Quero-Valenzuela F, Piedra-Fernández I, Hernández-Escobar F, Martínez-Ceres M, Romero-Palacios PJ, De Guevara ACL et al. Half the deaths after surgery for lung cancer occur after discharge. Surg Oncol 2018;27:630–4.
4. Hu Y, McMurry TL, Wells KM, Isbell JM, Stukenborg GJ, Kozower BD. Postoperative mortality is an inadequate quality indicator for lung cancer resection. Ann Thorac Surg 2014;97:973–9.
5. Tsukazan MTR, Terra RM, Vigo Á, Fortunato GA, Camargo SM, de Oliveira HA et al. Video-assisted thoracoscopic surgery yields better outcomes than thoracotomy for anatomical lung resection in Brazil: a propensity score-matching analysis using the Brazilian Society of Thoracic Surgery database. Eur J Cardiothorac Surg 2018;53:993–8.
6. Al-Ameri M, Bergman P, Franco-Cereceda A, Sartipy U. Video-assisted thoracoscopic versus open thoracotomy lobectomy: a Swedish nationwide cohort study. J Thorac Dis 2018;10:3499–506.
7. Dziedzic R, Marjanski T, Binczyk F, Polanska J, Sawicka W, Rzyman W. Favourable outcomes in patients with early-stage non-small-cell lung cancer operated on by video-assisted thoracoscopic surgery: a propensity score-matched analysis. Eur J Cardiothorac Surg 2018;54:547–53.
8. Falcoz PE, Puyraveau M, Thomas PA, Decaluwe H, Hürtgen M, Petersen RH et al. Video-assisted thoracoscopic surgery versus open lobectomy for primary non-small-cell lung cancer: a propensity-matched analysis of outcome from the European Society of Thoracic Surgeon database. Eur J Cardiothorac Surg 2016;49:602–9.
9. Flores RM, Park BJ, Dycoco J, Aronova A, Hirth Y, Rizk NP et al. Lobectomy by video-assisted thoracic surgery (VATS) versus thoracotomy for lung cancer. J Thorac Cardiovasc Surg 2009;138:11–8.
10. Pagès PB, Delpy JP, Orsini B, Gossot D, Baste JM, Thomas P et al.; Epithor Project French Society of Thoracic and Cardiovascular Surgery. Propensity score analysis comparing videothoracoscopic lobectomy with thoracotomy: a French Nationwide Study. Ann Thorac Surg 2016;101:1370–8.

11. Decaluwe H, Petersen RH, Hansen H, Piwkowski C, Augustin F, Brunelli A et al.; ESTS Minimally Invasive Thoracic Surgery Interest Group (MITIG). Major intraoperative complications during video-assisted thoracoscopic anatomical lung resections: an intention-to-treat analysis. Eur J Cardiothorac Surg 2015;48:588–99.

12. Embun R, Royo-Crespo I, Recuero Díaz JL, Bolufer S, Call S, Congregado M et al. Spanish Video-Assisted Thoracic Surgery Group: method, auditing, and initial results from a national prospective cohort of patients receiving anatomical lung resections. Arch Bronconeumol 2020;56:718–24.

13. Benedetto U, Head SJ, Angelini GD, Blackstone EH. Statistical primer: propensity score matching and its alternatives. Eur J Cardiothorac Surg 2018;53:1112–7.

14. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med 2009 10;28:3083–107.

15. Abdelsattar ZM, Allen MS, Shen KR, Cassivi SD, Nichols FC, Wigle DA et al. Variation in hospital adoption rates of video-assisted thoracoscopic lobectomy for lung cancer and the effect on outcomes. Ann Thorac Surg 2017;103:454–60.

16. Austin PC. Propensity-score matching in the cardiovascular surgery literature from 2004 to 2006: a systematic review and suggestions for improvement. J Thorac Cardiovasc Surg 2007;134:1128–35.

17. Yan TD, Black D, Bannon PG, McCaughan BC. Systematic review and meta-analysis of randomized and nonrandomized trials on safety and efficacy of video-assisted thoracic surgery lobectomy for early-stage non-small-cell lung cancer. J Clin Oncol 2009;27:2553–62.

18. Puri V, Gaissert HA, Wormuth DW, Grogan EL, Burfeind WR, Chang AC et al. Thoracic surgeons participants performing thoracoscopic lobectomy. Ann Thorac Surg 2019;107:202–8.

19. Paul S, Altorki NK, Sheng S, Lee PC, Harpole DH, Onaitis MW et al. Thoracoscopic lobectomy is associated with lower morbidity than open lobectomy: a propensity-matched analysis from the STS database. J Thorac Cardiovasc Surg 2010;139:366–78.

20. Bongiolatti S, Gonfiotti A, Viggiano D, Borgianni S, Politi L, Crisci R et al.; Italian VATS Group. Risk factors and impact of conversion from VATS to open lobectomy: analysis from a national database. Surg Endosc 2019 1;33:3953–62.

21. Laursen LØ, Petersen RH, Hansen HJ, Jensen TK, Ravn J, Konge L. Video-assisted thoracoscopic surgery lobectomy for lung cancer is associated with a lower 30-day morbidity compared with lobectomy by thoracotomy. Eur J Cardiothorac Surg 2016;49:870–5.

22. Power AD, Merritt RE, Abdel-Rasoul M, Moffatt-Bruce SD, D'Souza DM, Kneuertz PJ. Estimating the risk of conversion from video-assisted thoracoscopic lung surgery to thoracotomy—a systematic review and meta-analysis. J Thorac Dis 2021;13:812–23.

23. Fourdrain A, De Dominicis F, Iquille J, Lafitte S, Merlusca G, Witte-Pfister A et al. Intraoperative conversion during video-assisted thoracoscopy does not constitute a treatment failure. Eur J Cardiothorac Surg 2019;55:660–5.

24. Brunelli A, Dinesh P, Woodcock-Shaw J, Littlechild D, Pompili C. Ninety-day mortality after video-assisted thoracoscopic lobectomy: incidence and risk factors. Ann Thorac Surg 2017;104:1020–6.

25. Gómez de Antonio D, Crowley Carrasco S, Romero Román A, Royuela A, Sánchez Calle Á, Obiols Fornell C et al. Surgical risk following anatomic lung resection in thoracic surgery: a prediction model derived from a Spanish Multicenter Database. Arch Bronconeumol 2021 Feb 24:S0300-2896(21)00070-3.

26. Aguinagalde B, Insausti A, Lopez I, Sanchez L, Bolufer S, Embun R. VATS lobectomy morbidity and mortality is lower in patients with the same ppoDLCO: analysis of the database of the Spanish Video-Assisted Thoracic Surgery Group. Arch Bronconeumol 2021;57:750–6.

27. Xu J, Ni H, Wu Y, Cao J, Han X, Liu L et al. Perioperative comparison of video-assisted thoracic surgery and open lobectomy for pT1-stage non-small cell lung cancer patients in China: a multi-center propensity score-matched analysis. Transl Lung Cancer Res 2021;10:402–14.

Figure_1

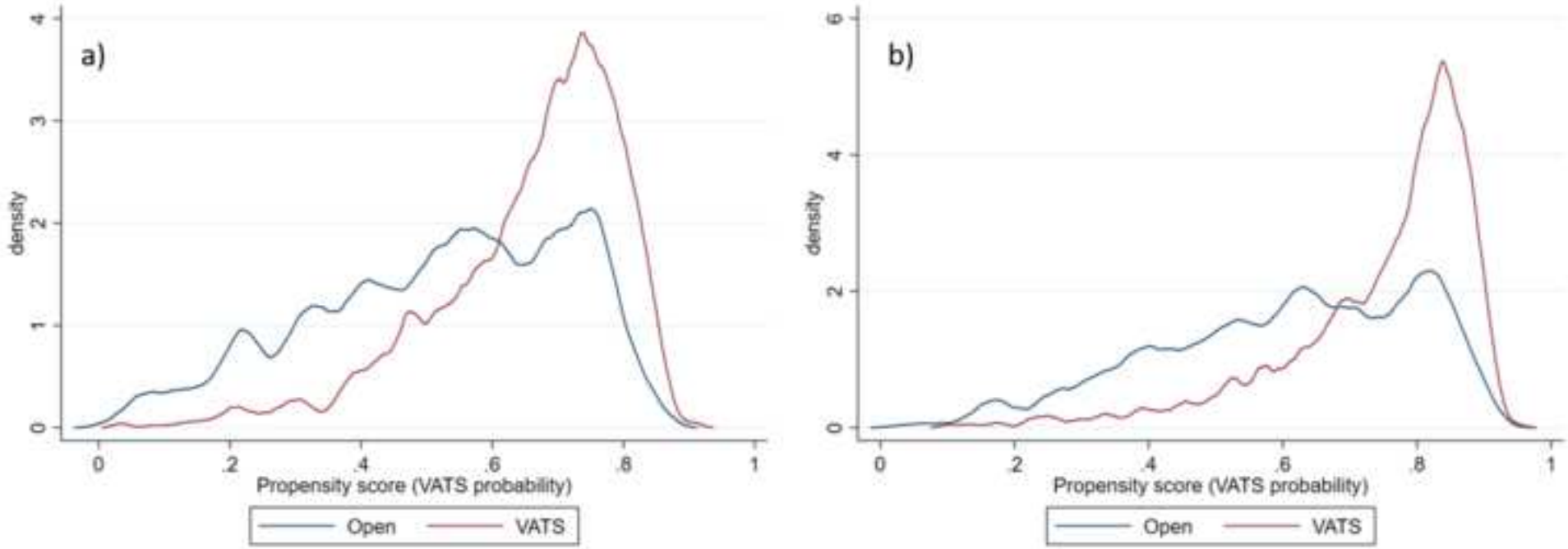

a)
b)
density
Propensity score (VATS probability)
0
.2
.4
.6
.8
1
Open
VATS

Figure_2

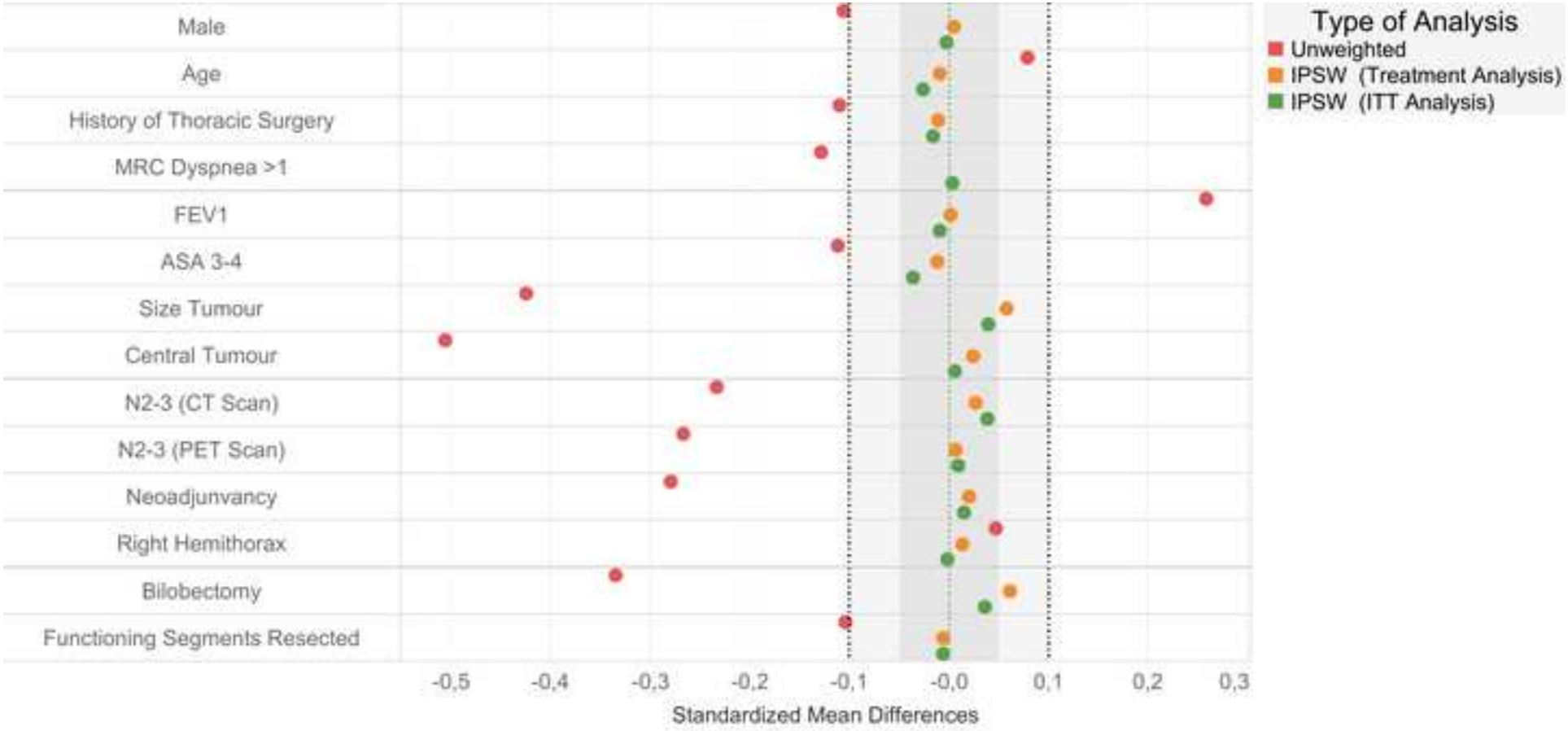

Male
Age
History of Thoracic Surgery
MRC Dyspnea >1
FEV1
ASA 3-4
Size Tumour
Central Tumour
N2-3 (CT Scan)
N2-3 (PET Scan)
Neoadjunvancy
Right Hemithorax
Bilobectomy
Functioning Segments Resected
-0,5
-0,4
-0,3
-0,2
-0,1
-0,0
0,1
0,2
0,3
Standardized Mean Differences
Type of Analysis
Unweighted
IPSW (Treatment Analysis)
IPSW (ITT Analysis)

Figure_3

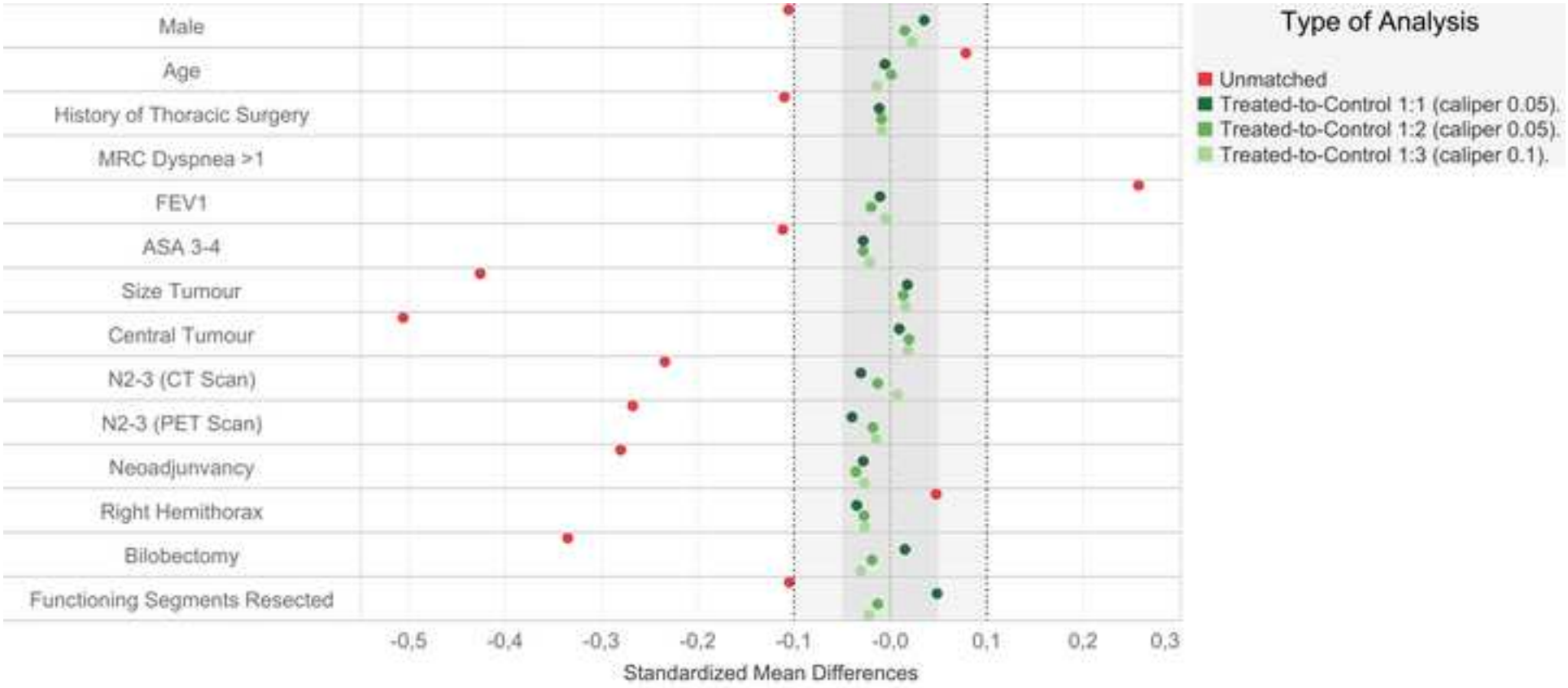
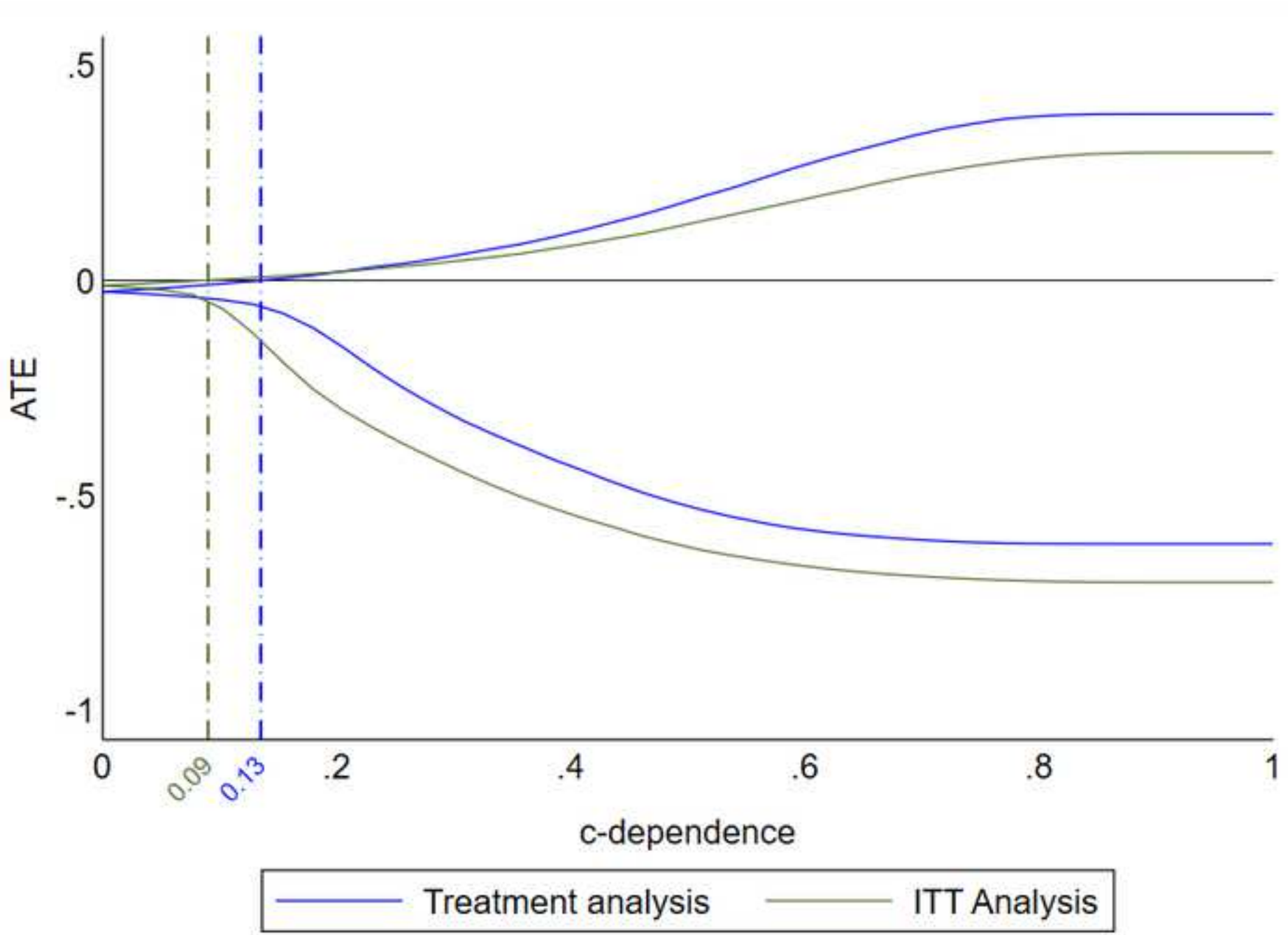

Male
Age
History of Thoracic Surgery
MRC Dyspnea >1
FEV1
ASA 3-4
Size Tumour
Central Tumour
N2-3 (CT Scan)
N2-3 (PET Scan)
Neoadjunvancy
Right Hemithorax
Bilobectomy
Functioning Segments Resected
-0,5
-0,4
-0,3
-0,2
-0,1
-0,0
0,1
0,2
0,3
Standardized Mean Differences
Type of Analysis
Unmatched
Treated-to-Control 1:1 (caliper 0.05).
Treated-to-Control 1:2 (caliper 0.05).
Treated-to-Control 1:3 (caliper 0.1).

Figure_4


.5
0
-.5
-1
ATE
0
0.09
0.13
.2
.4
.6
.8
1
c-dependence
Treatment analysis
ITT Analysis