The fundamental problem of confounding by medical operability in retrospective comparisons of surgery versus stereotactic body radiation therapy for early-stage lung cancer

Vivek Verma¹, Chad G. Rusthoven²

¹Department of Radiation Oncology, Allegheny General Hospital, Pittsburgh, PA, USA; ²Department of Radiation Oncology, University of Colorado School of Medicine, Aurora, CO, USA

Correspondence to: Chad G. Rusthoven, MD. Department of Radiation Oncology, University of Colorado School of Medicine, 1665 Aurora Court, Suite 1032, Aurora, CO 80045, USA. Email: chad.rusthoven@ucdenver.edu.

Provenance: This is an invited Editorial commissioned by the Section Editor Dr. Jie Dai (Department of Thoracic Surgery, Shanghai Pulmonary Hospital, Tongji University, Shanghai, China).

Comment on: Ajmani GS, Wang CH, Kim KW, *et al.* Surgical quality of wedge resection affects overall survival in patients with early stage non-small cell lung cancer. J Thorac Cardiovasc Surg 2018;156:380-91.e2.

Submitted Jun 05, 2018. Accepted for publication Jun 19, 2018. doi: 10.21037/jtd.2018.06.99 View this article at: http://dx.doi.org/10.21037/jtd.2018.06.99

The standard of care for early-stage non-small cell lung cancer (NSCLC) is lobectomy with lymph node (LN) dissection (1). Sublobar resection [e.g., wedge resection (WR) or anatomic segmentectomy] is an alternative for patients unsuitable for lobectomy, owing to a greater degree of lung preservation and potentially less postoperative morbidity and/or mortality (2). However, the broader candidacy of sublobar resection comes at the theoretical expense of a decrease in the "oncologic quality" of resection, which may lead to poorer tumor-related outcomes. This was documented in the classic Lung Cancer Study Group randomized study (albeit with antiquated surgical techniques and technology) illustrating poorer local control and cancer-specific survival with sublobar resection as compared to lobectomy (3). Retrospective studies with more contemporary surgical techniques have demonstrated somewhat conflicting results (4,5). Thus, the comparative effectiveness of lobar versus sublobar resection in carefully selected patients remains unresolved in the context of contemporary management.

Stereotactic body radiation therapy (SBRT) represents the preferred therapy for medically inoperable early-stage NSCLC (1), whereas the role of SBRT among medically operable patients is a topic of substantial ongoing research and debate. Interestingly, the only available randomized comparison of lobectomy versus SBRT to date demonstrated improved overall survival (OS) for the latter (6). However, that data has multiple caveats including the pooling of data from two poorly-accrued randomized trials, the resulting small sample sizes, and the underutilization of minimally-invasive surgical techniques such as videoassisted thoracoscopic surgery (VATS). Thus, achieving a better understanding of differences in outcomes following surgical resection versus SBRT in operable populations remains an important goal, for which existing retrospective analyses on the topic have offered conflicting results (7-9).

The article to which this editorial pertains was a retrospective analysis of the National Cancer Database (NCDB) performed by Ajmani and colleagues (10). Its goal was to evaluate the novel metric of WR "quality", defining high quality as resection with negative surgical margins and ≥ 5 LNs sampled, average quality as negative surgical margins with <5 LNs, and poor quality as positive surgical margins regardless of LN sampling. Using these three study-defined quality tiers, the authors reported improved OS with increasing surgical quality. The authors additionally performed an analysis to compare SBRT and WR, which illustrated that OS following SBRT was similar to low quality WR and inferior to high and average quality WR.

The authors of this study should be commended for their efforts to better characterize a novel correlate of OS in early-stage NSCLC treated with WR. However, similar to



Figure 1 OS for stage I NSCLC patients in retrospective comparisons of (A) "quality" of wedge resection *vs.* SBRT in a study by Ajmani *et al.* (10) and (B) SBRT in all patients, stratified by medically operable *vs.* inoperable status in a study by Onishi *et al.* (12). Both are presented with permission from the respective publishers. OS, overall survival; NSCLC, non-small cell lung cancer.

many other retrospective comparisons of surgical resection and SBRT, the analysis by Ajmani and colleagues suffers from the fundamental, uncontrolled confounding factor of medical operability (11). Medical operability is determined on the basis of several clinical factors including age, smoking history, performance status, specific comorbidities, and cardiopulmonary function. Each of these can substantially influence OS even when all patients receive the same therapy (12,13). To that end, a prior SBRT analysis reported 3-year OS rates of 77% and 43% for medically operable and inoperable patients, respectively (12), the former of which is numerically similar to the results of the recently reported Radiation Therapy Oncology Group 0618 trial of SBRT for medically operable patients (14). In turn, both are also similar to the 3-year OS observed by Ajmani et al. of approximately 77% and 53% in their medically operable high quality WR cohort and predominantly medically inoperable SBRT cohort, respectively (Figure 1). Moreover, a meta-analysis of stage I NSCLC patients treated with SBRT versus surgery revealed that OS differences became non-significant (and numerically favored SBRT) when models were adjusted for medical operability (15). Patients undergoing surgery in the analysis by Ajmani (10) and similar retrospective analyses are, by definition, medically operable; however, in accordance with national guidelines, the vast majority of patients offered SBRT outside of clinical trials are medically inoperable (1). This particular form of selection bias, known as indication bias, is recognized as a critical form of confounding in retrospective analyses that cannot be resolved by statistical adjustment (16).

The importance of baseline prognostic differences

between medically operable and inoperable patients cannot be overstated, and is particularly relevant to analyses of datasets such as the NCDB, which neither contain the vast majority of items used to define operability, nor measure non-OS endpoints. Beyond the complete absence of dedicated pulmonary and cardiac function data, it is also important to note that comorbidity scores are not synonymous with performance status (17). For example, based on NCDB coding, a patient with peptic ulcer and rheumatologic disease would have a higher Charlson-Deyo score (score of 2) than a patient with chronic pulmonary disease (score of 1). This disconnect likely explains why SBRT patients paradoxically had comorbidity scores superior to WR patients in the analysis by Ajmani (10), when performance status would almost certainly favor the surgical cohort based on the treatment recommendations of the national guidelines. The lack of this information also underscores the inability of propensity score matching and multivariate analyses to account for unrecorded factors. For instance, the propensity matched survival curve from Ajmani et al. (10) shows a precipitous drop in survival immediately after surgery in the low quality cohort (Figure 1). This finding, despite propensity matching, strongly suggests the presence of uncontrolled baseline differences in the low versus higher quality WR cohorts that may be more likely to drive early mortality than treatment-related factors such as margin status and extent of nodal sampling (which would be expected to preferentially impact long-term oncologic outcomes). Thus, it would be highly questionable to imply that positive surgical margins (the definition of low quality WR in that study) was the cause of the substantially

Study, year	Trial name	Three-year OS (%)
SBRT (operable)		
Chang <i>et al.</i> , 2015 (6)	Pooled analysis of STARS/ROSEL	95
Nagata e <i>t al.</i> , 2015 (13)	JCOG 0403	77
Shibamoto <i>et al.</i> , 2015 (18)	Japanese multicenter	73
Timmerman <i>et al.</i> , 2018 (14)	RTOG 0618	77
Surgical therapy*		
Chang <i>et al.</i> , 2015 (6)	Pooled analysis of STARS/ROSEL	79
Fernando <i>et al.</i> , 2014 (19)	ACOSOG Z4032	71
Su <i>et al.</i> , 2014 (20)	ACOSOG Z0030	76

Table 1 OS for early stage NSCLC patients in contemporary surgical trials as well as prospective trials of medically operable candidates receiving SBRT

*, Lobectomy utilized in two studies (6,20) with sublobar resection in a third study (19). OS, overall survival; NSCLC, non-small cell lung cancer; SBRT, stereotactic body radiation therapy; JCOG, Japan Clinical Oncology Group; RTOG, Radiation Therapy Oncology Group; ACOSOG, American College of Surgeons Oncology Group.

increased rates of immediate post-operative mortality in the low quality cohort, just as it would be misleading to suggest that high quality WR was the casual driver of the massive 24% OS advantage at 3 years over SBRT observed in this retrospective analysis.

In light of the fundamental confounding by operability and indication, prospective randomized trials likely represent the only suitable means to compare OS between SBRT and surgery for early-stage NSCLC. In the meantime, outcomes from prospective SBRT trials in medically operable patients (6,13,14,18) have demonstrated 3-year OS rates in the range of 73% to 95% (Table 1), which compare well with those of contemporary high-quality ACOSOG prospective surgical trials (71–76%) (6,19,20), as well as the 77% rate in the high quality WR cohort of the current NCDB study (10). Data from these prospective trials represent a far superior level of evidence for characterizing SBRT outcomes in medically operable cohorts as compared to retrospective analyses from datasets like the NCDB, which are more appropriate for characterizing observational outcomes with SBRT in predominantly inoperable populations. Importantly, these prospective results also provide sufficient justification for enrollment onto ongoing randomized trials (NCT02468024, NCT01753414, NCT02984761, NCT02629458).

With respect to the novel metric of surgical quality presented in this analysis, there are some additional caveats worth considering. First, using the association of surgical margins with OS as an indicator of surgical quality may be subject to confounding from tumor

biology, because aggressive growth patterns may cause more local parenchymal and lymphangitic involvement than appreciated on preoperative imaging, resulting in potentially higher rates of positive margins in more biologically aggressive tumors. Furthermore, it may not be accurate to categorically align the numerical extent of nodal sampling with surgical quality, as various anatomic factors (e.g., central vs. peripheral tumor location) can impact both the technical capability of performing WR as well as rates of occult nodal involvement. The NCDB's lack of information on sampled nodal stations is also noteworthy, as dissecting seven N1 LNs, for example, may be qualitatively different than removing the same number from both N1 and N2 stations, particularly since lobar location can impact nodal drainage and failure patterns (21). Selection bias is also an important concern, as healthier patients tend to be offered more aggressive oncologic therapies, including more extensive LN dissections. Finally, stage migration should also be considered, as increasing the number of LNs sampled without identifying metastases also increases the probability that a patient is truly N0, whereas patients with fewer or no LNs sampled may be more likely to have a false-negative N0 status. Occult node-positive cases will certainly exhibit a worse baseline prognosis than true N0 cases due to tumor biology, but subsequent treatmentrelated considerations are also relevant because patients with occult nodal involvement are unlikely to receive survivalextending adjuvant therapies appropriate for node-positive disease. Despite numerous retrospective analyses suggesting

Journal of Thoracic Disease, Vol 10, Suppl 18 July 2018

survival advantages with more extensive nodal sampling, randomized trials across disease sites (including lung cancer) (22-24) have often failed to confirm OS benefits; thus, causative relationships between increased LN sampling and OS remains unlikely. Indeed, the false appearance of improved survival associated with stage migration following more sensitive staging techniques (including extensive LN dissections), termed the "Will Rogers phenomenon" (25), is well characterized in the oncologic literature.

In summary, the analysis by Ajmani and colleagues presents observational outcomes following WR for earlystage NSCLC, focusing on the proposed metric of WR quality. Their hypothesis-generating analysis of treatment quality based on surgical margin status and extent of LN sampling may warrant further investigation, but may also suffer analytically from potential confounding in the areas of selection bias, tumor biology, technical factors including anatomic tumor location, and stage migration with more extensive LN sampling. The authors also compared the OS outcomes of their medically operable WR cohort with an SBRT cohort comprised of (as dictated by the national guidelines) primarily medically inoperable patients. Unfortunately, the comparison of prognostically distinct medically operable and predominantly inoperable cohorts suffers from fundamental, uncontrolled confounding in this NCDB dataset and many similar retrospective analyses. While further prospective randomized data comparing SBRT and surgery are awaited, the available prospective data (Table 1) represent the highest level of evidence for characterizations of SBRT outcomes in medically operable early-stage NSCLC populations. Overall, a greater awareness and acknowledgement of confounding by operability is important to both the interpretation of the early-stage NSCLC literature and to shared decisionmaking discussions between providers and patients.

Acknowledgements

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. National Comprehensive Cancer Network. Non-Small

Cell Lung Cancer. Version 4.2018. Available online: https://www.nccn.org/professionals/physician_gls/pdf/ nscl.pdf

- Stokes WA, Bronsert MR, Meguid RA, et al. Post-Treatment Mortality After Surgery and Stereotactic Body Radiotherapy for Early-Stage Non-Small-Cell Lung Cancer. J Clin Oncol 2018;36:642-51.
- Ginsberg RJ, Rubinstein LV. Randomized trial of lobectomy versus limited resection for T1 N0 non-small cell lung cancer. Lung Cancer Study Group. Ann Thorac Surg 1995;60:615-22; discussion 622-3.
- Zhang Y, Sun Y, Wang R, et al. Meta-analysis of lobectomy, segmentectomy, and wedge resection for stage I non-small cell lung cancer. J Surg Oncol 2015;111:334-40.
- Bedetti B, Bertolaccini L, Rocco R, et al. Segmentectomy versus lobectomy for stage I non-small cell lung cancer: a systematic review and meta-analysis. J Thorac Dis 2017;9:1615-23.
- Chang JY, Senan S, Paul MA, et al. Stereotactic ablative radiotherapy versus lobectomy for operable stage I nonsmall-cell lung cancer: a pooled analysis of two randomised trials. Lancet Oncol 2015;16:630-7. Erratum in: Lancet Oncol 2015;16:e427.
- Wang HH, Zhang CZ, Zhang BL, et al. Sublobar resection is associated with improved outcomes over radiotherapy in the management of high-risk elderly patients with Stage I non-small cell lung cancer: a systematic review and metaanalysis. Oncotarget 2017;8:6033-42.
- Chen H, Laba JM, Boldt RG, et al. Stereotactic Ablative Radiation Therapy Versus Surgery in Early Lung Cancer: A Meta-analysis of Propensity Score Studies. Int J Radiat Oncol Biol Phys 2018;101:186-94.
- Grills IS, Mangona VS, Welsh R, et al. Outcomes after stereotactic lung radiotherapy or wedge resection for stage I non-small-cell lung cancer. J Clin Oncol 2010;28:928-35.
- Ajmani GS, Wang CH, Kim KW, et al. Surgical quality of wedge resection affects overall survival in patients with early stage non-small cell lung cancer. J Thorac Cardiovasc Surg 2018;156:380-91.e2.
- 11. Stokes WA, Rusthoven CG. Surgery vs. SBRT in retrospective analyses: confounding by operability is the elephant in the room. J Thorac Dis 2018;10:S2007-10.
- Onishi H, Shirato H, Nagata Y, et al. Hypofractionated stereotactic radiotherapy (HypoFXSRT) for stage I nonsmall cell lung cancer: updated results of 257 patients in a Japanese multi-institutional study. J Thorac Oncol 2007;2:S94-100.
- 13. Nagata Y, Hiraoka M, Shibata T, et al. Prospective Trial of

Stereotactic Body Radiation Therapy for Both Operable and Inoperable T1N0M0 Non-Small Cell Lung Cancer: Japan Clinical Oncology Group Study JCOG0403. Int J Radiat Oncol Biol Phys 2015;93:989-96.

- Timmerman RD, Paulus R, Pass HI, et al. Stereotactic Body Radiation Therapy for Operable Early-Stage Lung Cancer: Findings From the NRG Oncology RTOG 0618 Trial. JAMA Oncol 2018. [Epub ahead of print].
- Zheng X, Schipper M, Kidwell K, et al. Survival outcome after stereotactic body radiation therapy and surgery for stage I non-small cell lung cancer: a meta-analysis. Int J Radiat Oncol Biol Phys 2014;90:603-11.
- Bosco JL, Silliman RA, Thwin SS, et al. A most stubborn bias: no adjustment method fully resolves confounding by indication in observational studies. J Clin Epidemiol 2010;63:64-74.
- Firat S, Bousamra M, Gore E, et al. Comorbidity and KPS are independent prognostic factors in stage I nonsmall-cell lung cancer. Int J Radiat Oncol Biol Phys 2002;52:1047-57.
- Shibamoto Y, Hashizume C, Baba F, et al. Stereotactic body radiotherapy using a radiobiology-based regimen for stage I non-small-cell lung cancer: five-year mature results. J Thorac Oncol 2015;10:960-4.
- Fernando HC, Landreneau RJ, Mandrekar SJ, et al. Impact of brachytherapy on local recurrence rates after sublobar resection: results from ACOSOG Z4032 (Alliance), a phase III randomized trial for high-risk operable nonsmall-cell lung cancer. J Clin Oncol 2014;32:2456-62.
- 20. Su S, Scott WJ, Allen MS, et al. Patterns of survival and recurrence after surgical treatment of early stage non-small

Cite this article as: Verma V, Rusthoven CG. The fundamental problem of confounding by medical operability in retrospective comparisons of surgery versus stereotactic body radiation therapy for early-stage lung cancer. J Thorac Dis 2018;10(Suppl 18):S2176-S2180. doi: 10.21037/jtd.2018.06.99

cell lung carcinoma in the ACOSOG Z0030 (ALLIANCE) trial. J Thorac Cardiovasc Surg 2014;147:747-52: Discussion 752-3. Erratum in: J Thorac Cardiovasc Surg 2014;148:756.

- Cerfolio RJ, Bryant AS. Distribution and likelihood of lymph node metastasis based on the lobar location of nonsmall-cell lung cancer. Ann Thorac Surg 2006;81:1969-73; discussion 1973.
- 22. Darling GE, Allen MS, Decker PA, et al. Randomized trial of mediastinal lymph node sampling versus complete lymphadenectomy during pulmonary resection in the patient with N0 or N1 (less than hilar) non-small cell carcinoma: results of the American College of Surgery Oncology Group Z0030 Trial. J Thorac Cardiovasc Surg 2011;141:662-70.
- 23. Giuliano AE, Ballman KV, McCall L, et al. Effect of Axillary Dissection vs No Axillary Dissection on 10-Year Overall Survival Among Women With Invasive Breast Cancer and Sentinel Node Metastasis: The ACOSOG Z0011 (Alliance) Randomized Clinical Trial. JAMA 2017;318:918-26.
- 24. ASTEC study group, Kitchener H, Swart AM, et al. Efficacy of systematic pelvic lymphadenectomy in endometrial cancer (MRC ASTEC trial): a randomised study. Lancet 2009;373:125-36. Erratum in: Lancet 2009;373:1764.
- 25. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. N Engl J Med 1985;312:1604-8.

S2180