Analysis of longitudinal quality-of-life data in high-risk operable patients with lung cancer: Results from the ACOSOG Z4032 (Alliance) multicenter randomized trial

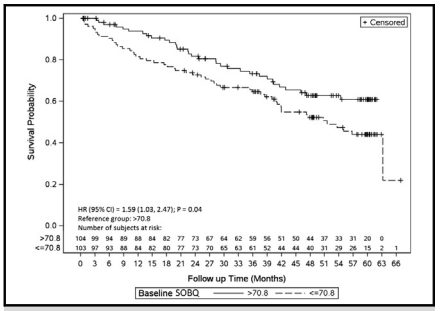
Hiran C. Fernando, MD,[a] Rodney J. Landreneau, MD,[b] Sumithra J. Mandrekar, PhD,[c] Francis C. Nichols, MD,[d] Thomas A. DiPetrillo, MD,[e] Bryan F. Meyers, MD,[f] Dwight E. Heron, MD,[b] Shauna L. Hillman, MS,[c] David R. Jones, MD,[g] Sandra L. Starnes, MD,[h] Angelina D. Tan, BS,[c] Benedict D. T. Daly, MD,[b] and Joe B. Putnam, MD,[i] for the Alliance for Clinical Trials in Oncology

ABSTRACT

Background: Prior studies have suggested that low baseline quality-of-life (QOL) scores predict worse survival in patients undergoing lung cancer surgery. However, these studies involved average-risk patients undergoing lobectomy. We report QOL results from a multicenter trial, American College of Surgeons Oncology Group Z4032, which randomized high-risk operable patients to sublobar resection (SR), or SR with brachytherapy, and included longitudinal QOL assessments.

Methods: Global QOL, using the 36-item Short-Form Health Survey (SF36), and the dyspnea score from the University of California, San Diego Shortness of Breath Questionnaire (SOBQ) scale, was measured at baseline, 3, 12, and 24 months. SF36 physical component summary (PCS) and mental component summary (MCS) scores were standardized and adjusted for age and gender normals, with scores <50 indicating below-average health status. SOBQ scores were transformed to a 0-100 (poor-excellent) scale. Aims were to: (1) determine the impact of baseline scores on recurrence-free survival, overall survival, and 30-day adverse events (AEs); and (2) identify subgroups (surgical approach, resection type. tumor location, tumor size, respiratory function) with a ≥10-point decline or improvement in QOL after SR.

Results: Two hundred twelve eligible patients were included. There were no significant differences in baseline QOL scores between arms. Median baseline PCS, MCS, and SOBQ scores were 42.7, 51.1, and 70.8, respectively. There were no differences in grade-3+ AEs, overall survival, or recurrence-free survival in patients with baseline scores ≤median versus >median values, except for a significantly worse overall survival for patients with baseline SOBQ scores ≤median value. There were no significant differences between the study arms in percentage change of QOL scores from baseline to 3, 12, or 24 months. Further comparison combining the 2 arms demonstrated a higher percentage of patients with a ≥10-point decline in SOBQ scores with segmentectomy compared with wedge



Overall survival after sublobar resection, comparing patients with low and high baseline dyspnea scores.

Central Message

Although low dyspnea scores were associated with worse survival after surgery in high-risk patients with lung cancer, poor global quality-of-life scores had none.

Author Perspective

Serial quality of life (QOL) is reported in high-risk operable patients undergoing lung resection. Low global QOL had no impact, although low baseline dyspnea scores were associated with worse long-term survival. VATS was associated with a more rapid return to baseline function, suggesting preferential use of VATS for this patient population.

From the Boston Medical Center,[a] Boston, Mass; University of Pittsburgh,[b] Pittsburgh, Pa; Alliance Statistics and Data Center,[c] Mayo Clinic, Rochester, Minn; Mayo Clinic,[d] Rochester, Minn; Rhode Island Hospital,[e] Providence, RI; Washington University of St Louis,[f] St Louis, Mo; University of Virginia,[g] Charlottesville, Va; University of Cincinnati,[h] Cincinnati, Ohio; and Vanderbilt University Medical Center,[i] Nashville, Tenn.

Supported by the National Cancer Institute (Bethesda, Md) (CA076001), and by a grant from Oncura (Princeton, NJ). The study was supported, in part, by grants from the National Cancer Institute (Bethesda, Md) (CA31946) to the Alliance for Clinical Trials in Oncology (Monica M. Bertagnolli, MD, Chair) and to the Alliance Statistics and Data Center at Mayo Clinic (Rochester, Minn) (to Daniel J. Sargent, PhD, CA33601). (ClinicalTrials.gov number NCT00107172).

Read at the 94th Annual Meeting of The American Association for Thoracic Surgery, Toronto, Ontario, Canada, April 26-30, 2014.

Received for publication April 4, 2014; revisions received Sept 10, 2014; accepted for publication Nov 4, 2014; available ahead of print Dec 10, 2014.

Address for reprints: Hiran C. Fernando, MD, Department of Cardiothoracic Surgery, Boston Medical Center, 88 East Newton St, Boston, MA 02118 (E-mail: Hiran.Fernando@bmc.org).

0022-5223/$36.00

http://dx.doi.org/10.1016/j.jtcvs.2014.11.003

Abbreviations and Acronyms

AE = adverse event
DLCO = diffusion capacity of the lung for carbon monoxide
FEV1 = forced expiratory volume in 1 second
MCS = mental component summary
PCS = physical component summary
QOL = quality of life
SF36 = 36-item short-form health survey
SOBQ = shortness of breath questionnaire
SR = sublobar resection
SRB = sublobar resection with brachytherapy
VATS = video-assisted thoracic surgery

resection (40.5% vs 21.9%, $P = .03$) at 12 months, with thoracotomy versus video-assisted thoracic surgery (VATS) (38.8% vs 20.4%, $P = .03$) at 12 months, and T1b versus T1a tumors (46.9% vs 23.5%, $P = .020$) at 24 months. A ≥10-point improvement in PCS score was seen at 3 months with VATS versus thoracotomy (16.5% vs 3.6%, $P = .02$).

Conclusions: In high-risk operable patients, poor baseline QOL scores were not predictive for worse overall or recurrence-free survival, or for higher risk for AEs following SR. VATS was associated with improvement in physical function at 3 months, and improved dyspnea scores at 12 months, lending support for the preferential use of VATS when SR is undertaken. (J Thorac Cardiovasc Surg 2015;149:718-26)

Sublobar resection (SR) has traditionally been used for high-risk operable patients with non–small cell lung cancer when lobectomy is not considered feasible. More recently, nonoperative treatments, such as stereotactic body radiation therapy and radiofrequency ablation have been applied to this population, after successful application in medically inoperable patients.[1,2] Standard outcome measures, such as survival and recurrence rates, are undoubtedly the most helpful measures to guide physicians in making treatment recommendations. Quality of life (QOL), however, is an important variable that rarely has been measured in these clinical trials, but it is of tremendous significance, particularly when treating high-risk operable patients, who often have emphysema and early-stage lung cancer.

The American College of Surgeons Oncology Group Z4032 (Alliance) was a randomized trial undertaken to compare SR alone with SR with brachytherapy (SRB) for high-risk operable patients with early-stage non–small cell lung cancer. The primary endpoint was the time to local recurrence, comparing these 2 arms that utilized only SR in this high-risk operable population. The difference was not significant and has been reported elsewhere.[3] A secondary aim of this study was to measure longitudinal QOL, and self-reported functional health status. We report these outcomes in this article.

METHODS

Eligible patients for this study included patients with biopsy-proven stage-I lung cancers ≤3 cm in maximum diameter. Patients were defined as high risk for lobectomy if they met ≥1 major criterion or 2 minor criteria.[3] Patients were required to be evaluated by a thoracic surgeon approved by the American College of Surgeons Oncology Group. In addition, patients had to be considered either to not be a candidate for lobectomy (standard-risk operable patients), or to not be a candidate for any form of pulmonary resection (medically inoperable patients). To confirm that patients did not have nodal involvement, all suspicious lymph nodes seen on positron emission tomography or computed tomography scan required biopsy by mediastinoscopy, endobronchial ultrasound, or lymph node sampling at the time of resection. Wedge or segmental resection was allowed, and could be performed by video-assisted thoracic surgery (VATS) or thoracotomy.

Two methods of brachytherapy were allowed.[4,5] In the first technique, polyglactin sutures containing ^{125}I seeds (Oncura, Princeton, NJ) were placed parallel to and 5 mm away from the staple line on each side of the resection margin. The suture strands were fixed to the lung surface with several 3.0 silk or polyglactin sutures placed 1-2 cm apart. With the second brachytherapy technique, a polyglycolic mesh implant was created. The same ^{125}I suture strands were woven into a piece of Vicryl (polyglactin 910; Ethicon, Inc, Somerville, NJ) mesh. The strands were placed at 1-cm intervals. The mesh was then sutured over the staple line. The dosimetry goal of the brachytherapy was to deliver 100 Gy at 5-7 mm along the central axis of the resection margin.

Adverse events (AEs) were recorded using the Common Terminology Criteria for Adverse Events,[6] version 3.0. A report of 30- and 90-day AEs has previously been published from this study.[7] No significant differences were found between the study arms in grade-3 or higher AEs.

Global QOL was measured using the 36-item short-form health survey (SF36), an instrument that has been reported and validated previously,[8] and provides a measure of overall health status. Scores can be reported as 8 domains of functional health and well-being, or transformed into a physical component summary (PCS) score and a mental component summary (MCS) score. In this study, SF36 results were expressed as PCS and MCS scores. Dyspnea was evaluated using the University of California, San Diego Shortness of Breath Questionnaire (SOBQ). This self-reported instrument that measures functional health status has been validated in other studies.[9] This is a 24-item disease-specific questionnaire that assesses self-reported shortness of breath experienced while performing activities of daily living. QOL assessments, using the SF36, and functional health status assessed using the SOBQ, were administered at baseline and at 3, 12, and 24 months.

All patients provided written informed consent before trial enrollment, in accordance with applicable guidelines. At each participating site, institutional review board approval was obtained in accordance with an assurance filed with and approved by the US Department of Health and Human Services. Data collection and statistical analyses were conducted by the Alliance Statistics and Data Center at Mayo Clinic (Rochester, Minn).

Statistical Analysis

All randomized and eligible patients are included in the QOL analysis. The SOBQ scores were converted to a percentage of theoretical range 0-100, with 0 = poor, and 100 = excellent. Eight subscale scores of SF36 were calculated by adding the subscale-related individual items and transforming them to 0-100, with 0 = poor, and 100 = excellent. Standardized scores of SF36 PCS and MCS scores were calculated using the mean, SD, and scoring coefficients from the US general population.

GTS

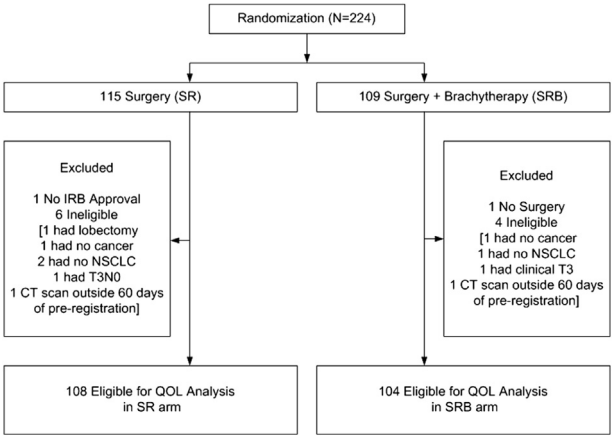


FIGURE 1. Patient CONSORT diagram. *SR*, Sublobar resection; *SRB*, SR with brachytherapy; *IRB*, institutional review board; *NSCLC*, non–small cell lung cancer; *T3N0*, tumor stage 3, nodal involvement; *T3*, tumor stage 3; *CT*, computed tomography; *QOL*, quality of life.

The SF36 PCS and SF36 MCS standardized scores were then adjusted for age and gender using the mean and SD of the US general population according to age and gender grouping, and employing a linear transformation.[9,10] Scores <50 indicate below-average health status. Compliance for SF36 and SOBQ at each time point of assessment was defined as the percentage of eligible patients who filled out the questionnaire (any item on the SF36 and SOBQ) among all evaluable (still on treatment) patients.

A clinically significant decline (improvement) in QOL was defined as a ≥10-point decrease (increase) from baseline.[11] In addition to considering the scores on a continuous scale, scores were dichotomized using the sample median (≤median vs >median). Scores at baseline, and at 3, 12, and 24 months, as well as percentage change in scores from baseline to 3, 12, and 24 months, were compared between the arms, using a Wilcoxon rank sum test. Baseline scores ≤median versus those >median were compared for patients who had any grade-3 or above AE within 30 days, versus those who had none, using the Fisher exact test. The 10-point decline in scores between the SR and the SRB arms from baseline to 3, 12, and 24 months were compared using a Fisher exact test. Similar analyses were carried out regardless of arm by: (1) resection type (wedge vs segmentectomy); (2) surgical approach (VATS vs thoracotomy); (3) clinical tumor size (≤2 cm vs >2 cm); (4) lobe (upper vs other); (5) any grade-3 or higher AE within 30 days (yes vs no); (6) median of baseline DLCO (diffusion capacity of the lung for carbon monoxide) (>45% vs ≤45%); and (7) median of baseline FEV1 (forced expiratory volume in 1 second) (>49% vs ≤49%), to identify potentially vulnerable or successful subgroups. Finally, a generalized estimating equation model was utilized to assess the impact of intervention arm and other baseline factors on longitudinal PCS, MCS, and SOBQ scores.

Overall survival was defined as the time from randomization to death due to any cause. Recurrence-free survival was defined as the time from randomization to the first of any recurrence or death from any cause. The distribution of survival times was estimated using the Kaplan-Meier method, and Cox proportional hazards models (adjusted and unadjusted for treatment arm) were used to evaluate the prognostic importance of baseline PCS, MCS, and SOBQ (as both continuous and categorized at the median value) on overall survival and recurrence-free survival. A landmark analysis at 3 and 12 months was also utilized to assess the impact of a 10-point decline in QOL scores from baseline to 3 months, and baseline to 12 months on subsequent overall survival and recurrence-free survival. Two-sided *P* values ≤.05 were considered statistically significant.

TABLE 1. Baseline patient characteristics

Factor	SR (N = 108)	SRB (N = 104)	*P* value*
Age (y; median, range)	70 (49-85)	71 (50-87)	.47†
Gender			.89
Female	61 (56.5)	57 (54.8)	
Male	47 (43.5)	47 (45.2)	
PS			.72
0	19 (17.6)	23 (22.1)	
1	63 (58.3)	58 (55.8)	
2	26 (24.1)	23 (22.1)	
Clinical nodule size			.78
≤2 cm	70 (64.8)	65 (62.5)	
>2 cm	38 (35.2)	39 (37.5)	
T stage			.12
T1	108 (100)	101 (97.1)	
T2	0 (0)	3 (2.9)	
T3	0 (0)	0 (0)	
M Stage: M0	108 (100)	104 (100)	NA
N Stage: N0	108 (100)	104 (100)	NA
ASA class on surgery day‡			.05
I/II	10 (9.3)	20 (19.2)	
III/IV	98 (90.7)	83 (79.8)	
Baseline FEV1 (%; median, range)‡	48 (22-117)	53 (25-110)	.31†
Baseline DLCO (%; median, range)‡	46 (18-97)	44 (8-83)	.25†

Values are n (%), unless otherwise indicated. *SR*, Sublobar resection; *SRB*, SR with intraoperative brachytherapy; *PS*, physical status; *T*, tumor; *M*, metastases; *N*, nodal involvement; *NA*, not applicable; *ASA*, American Society of Anesthesiologists; *FEV1*, forced expiratory volume in 1 second; *DLCO*, diffusion capacity of the lung for carbon monoxide. *Fisher exact test. †Wilcoxon rank sum test. ‡1 SRB patient with missing data.

RESULTS

Data were frozen for this analysis on July 15, 2013. A total of 224 patients were registered to the Z4032 trial. One patient from the SR arm had the intervention at a hospital that was not approved by an institutional review board, and was therefore deemed not evaluable. One patient randomized to the SRB arm did not have surgery and was also not evaluable. An additional 10 registered (6 SR and 4 SRB) patients were found to be ineligible (Figure 1). Thus, 212 patients (108 SR and 104 SRB) are included in this analysis. The completion rates for questionnaires at baseline, and at 3, 12, and 24 months for the SR and SRB arms, respectively, were 97.2% versus 99.0%, 82.4% versus 83.7%, 63.9% versus 74.0%, and 46.3% versus 53.8%. The drop in completion rates over time by study subjects may have been related to responder fatigue. Table 1 provides the baseline patient characteristics, by arm. Median length of follow-up on living patients was 4.4 years (range: 0.04-5.59 years).

The baseline median PCS, MCS, and SOBQ scores for the 212 patients were 42.7, 51.1, and 70.8, respectively. A total of 65% and 46.5% of patients in our study had baseline PCS and MCS scores, respectively, that were at least 1 SD below those of the US general population. Table 2 shows the standardized PCS and MCS scores based

TABLE 2. Mean scores of subscales comparison by age and gender between Z4032 (American College of Surgeons Oncology Group Z4032 multicenter trial) and normative data

		Physical component means*					Mental component means*				
Gender	Age (y)	Normative data (mean)	Z4032 data (mean)	Z4032 data (SD)	Z4032 data (N)	Difference in means (SD)	Normative data (mean)	Z4032 data (mean)	Z4032 data (SD)	Z4032 data (N)	Difference in means (SD)
Male	45-<55	50.40	28.68	2.82	3	2	51.03	45.36	2.32	3	0.5
Male	55-<65	46.90	38.02	11.18	21	1	51.60	50.99	8.32	21	About the same
Male	≥65	41.95	36.55	10.73	64	0.5	52.51	51.47	10.19	64	About the same
Female	45-<55	48.95	34.07	10.68	7	1.5	50.07	48.93	12.50	7	About the same
Female	55-<65	45.03	35.00	10.89	26	1	50.56	45.98	11.89	26	0.5
Female	≥65	41.02	34.88	9.47	83	0.5	51.44	51.19	10.40	83	About the same

Normative data are based on a 1990 US general population sample conducted by the National Opinion Research Center.[10] The PCS and MCS scores are standardized to a mean of 50 and a SD of 10. *SD*, Standard deviation. *Standardized scores were provided, which involved the following: (1) Z scores were calculated for each of the 8 subscales, utilizing the mean and SD of the US general population; (2) raw scores of physical and mental components were calculated by summation of the 8 subscales Z scores after multiplying each subscale score by the scoring coefficients from the US general population; and (3) standardized scores were calculated by multiplying the raw scores by 10, and then adding 50 to the scores.

on age and gender grouping. PCS scores were at least 0.5 to 2 SD lower than US general values for all groups. MCS scores were similar for most groups; in 2 groups, differences of 0.5 SD were seen.

Comparison of SRB and SR Arms

Median PCS, MCS, and SOBQ scores at each time point for each arm are depicted in Figure 2. There were no significant differences between arms at baseline, or at 3, 12, or 24 months. Based on the generalized estimating equation models, although PCS showed a significant trend over time ($P = .05$) and SOBQ ($P < .01$) scores, no significant differences by arm were observed for any of the scores (P for PCS = .74; P for MCS = .66; P for SOBQ = .48). The time trend was not significant when using data only up to 12 months. Additionally, the median % change in PCS, MCS, and SOBQ scores did not change significantly from baseline during follow-up. Therefore, the arms were combined for further analysis.

Longitudinal QOL and Functional Health Status for All Patients

At 12 months, there was a significantly greater proportion of patients with a ≥10-point decline in SOBQ scores among those who had segmental resection (40.5%) versus wedge resection (21.9%) ($P = .03$), and for patients who had a thoracotomy (38.8%) versus a VATS resection (20.4%) ($P = .03$). At 24 months, resection of tumors >2 cm (46.9%) was associated with a greater decline (of ≥10 points) in SOBQ scores than resection of tumors ≤2 cm (23.5%; $P = .02$). No subgroup was associated with any significant ≥10-point decline in PCS and MCS scores; however a ≥10-point improvement in PCS was seen at 3 months with VATS but not with thoracotomy (16.5% vs 3.6%, $P = .02$). There were no significant (≥10 points) declines in QOL scores at 3, 12, or 24 months of follow-up for patients with: upper versus lower lobe resections; any versus no grade-3 or greater AEs within 30 days; and baseline pulmonary function test scores below versus above median values.

Based on the generalized estimating equation models, a significant time trend for PCS ($P = .05$) and SOBQ ($P < .01$) scores was still observed. However, this trend was not significant when using data only up to 12 months. In addition, patients with baseline DLCO ≤45% had declining PCS ($P < .01$) and SOBQ ($P = .01$) scores over time.

Using baseline scores dichotomized at sample median (≤median vs >median), no differences were found in the occurrence of grade-3 or higher AEs at 30 days. Overall survival was significantly worse for patients with baseline SOBQ scores ≤median value (Figure 3). No significant differences were found in overall survival or recurrence-free survival by baseline PCS or MCS scores (either as continuous, or categoric at the median value). Table 3 shows the results of the landmark analysis for overall survival and recurrence-free survival for 10-point decline in SOBQ, PCS, and MCS scores from baseline to 3 and 12 months. Patients with a ≥10-point decline in SOBQ at 12 months had worse subsequent overall survival (hazard ratio, 2.10; 95% confidence interval: 1.16, 3.81; $P = .01$). None of the others was significantly associated with subsequent overall survival or recurrence-free survival.

DISCUSSION

A measure of QOL and functional health status is rarely reported in surgical publications, yet it is an important metric that can be of use to physicians and patients when making treatment decisions. Previous reports in the thoracic literature have usually involved standard-risk operable patients.[12,13] A recent study[13] of 245 patients treated with lobectomy or pneumonectomy measured QOL using the SF36. In that study, a PCS score of <50, as well as age >70 years and DLCO <70% were associated with poor overall survival. The patients in our study represented a high-risk operable group who were considered poor candidates for lobar

GTS

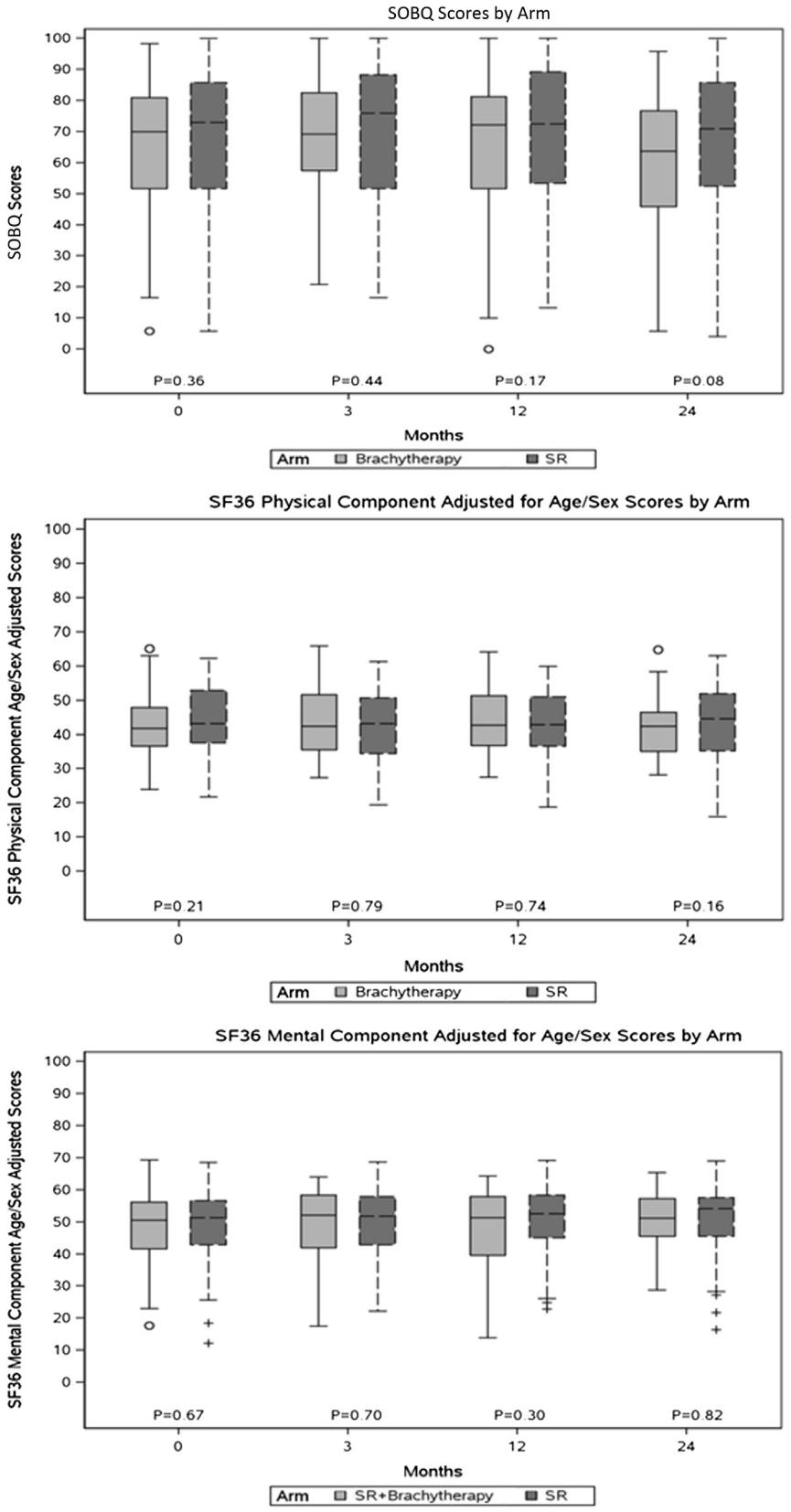


FIGURE 2. Comparison of quality-of-life scores at each time point, by arm. *SOBQ*, (University of California, San Diego) shortness of breath questionnaire; *SR*, sublobar resection; *SF36*, 36-item short-form health survey.

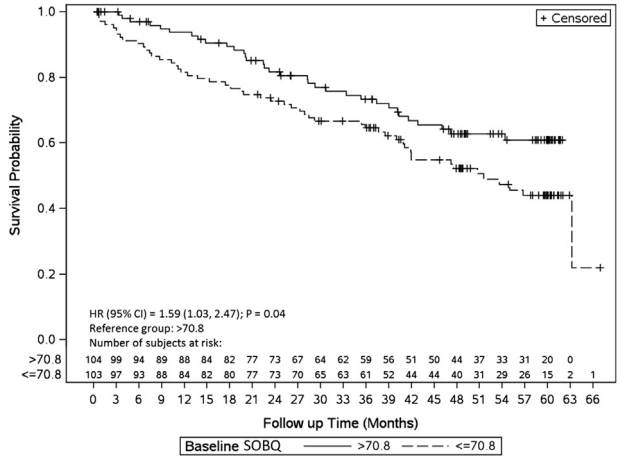


FIGURE 3. Kaplan-Meier curve for overall survival by median values of baseline SOBQ scores. *HR*, Hazard ratio; *SOBQ*, (University of California, San Diego) shortness of breath questionnaire; *CI*, confidence interval.

resection. It is noteworthy that the median ages of our patients were 70 and 71 years, and the median DLCOs were 46% and 44% for the SR and SRB arms, respectively.

Previous studies that measured QOL after lung resection have shown that thoracotomy is associated with a slower return to normal QOL compared with VATS.[12,14] In addition, studies have shown that more-complex resections, such as pneumonectomy, are associated with worse postoperative QOL.[14,15] Although none of our patients underwent such complex resections, the use of a segmental resection rather than wedge, as well as a thoracotomy rather than VATS, was associated with a larger proportion of patients with significant declines in SOBQ scores at the 12-month follow-up. In addition, in our series, VATS was associated with more-rapid improvement in PCS scores, as measured at 3 months after surgery. Although our results suggest that wedge resection may be preferential with respect to postoperative dyspnea, this approach has to be weighed against the oncologic benefits that have been reported with segmentectomy as opposed to wedge in other studies.[16]

Another finding from previous studies is that QOL will usually fall, at least with respect to physical functioning, in the early postoperative period, and then improve with time.[14,17] In one study, QOL decreased at 1 month but improved to baseline values by 3 months for lobectomy patients.[14] This improvement was not seen after pneumonectomy. In another study that involved 156 patients treated with lobectomy or pneumonectomy, SF36 was measured preoperatively, and at 1 and 3 months postoperatively.[17] PCS scores were significantly lower at 1 month compared with baseline. At 3 months, scores had recovered. MCS scores were unchanged. In our study, no significant decline was seen, but this may have been related to QOL being measured at 3 months rather than earlier after surgery.

We analyzed our data to see if low baseline QOL scores predicted poor survival. Although low baseline PCS and MCS scores did not predict poor survival, low baseline SOBQ scores did. As discussed earlier, previous studies in standard-risk operable patients have suggested that low PCS scores and DLCO are associated with poor overall survival.[13] A prospective study of 173 patients with clinical stage-I or stage-II non–small cell lung cancer measured QOL preoperatively and serially after surgery, for 2 years.[15] Recurrence occurred in 36% at 2 years. QOL improved in those patients without recurrence, whereas in patients with recurrence, some early recovery in QOL deteriorated significantly after 1 month. As discussed, our results suggest that in high-risk patients, baseline PCS and MCS scores are not good predictors of outcome.

The occurrence of postoperative complications could be postulated to predict poor QOL scores at longer follow-up. Certainly, this predictive value has been demonstrated in patients undergoing curative colorectal surgery.[18] In our study, no significant ($\geq$10 points) declines occurred in PCS, MCS, or SOBQ scores at 3, 12, or 24 months in patients either with or without grade-3 or higher AEs within 30 days. In addition, we performed a landmark analysis to determine whether postoperative scores could predict poor outcome. A significant decline in PCS, MCS, or SOBQ score at 3 months did not predict recurrence-free survival. However, a 10-point drop in SOBQ score at 12 months did predict poor subsequent overall survival.

Although QOL measurement may help surgeons decide what would be an optimal surgical therapy, it has even more relevance when considering surgical versus nonsurgical therapies for high-risk patients with early-stage lung cancer. A recent study of stereotactic body radiation therapy for medically inoperable lung cancer patients measured QOL before treatment, at 6 weeks, and serially for 12 months. The mean FEV1 and DLCO were 62.2% and 61.5%, respectively. A $\geq$10-point change was considered significant. QOL measurements that included SOBQ scores showed no significant decline. However, the mean DLCO did drop significantly after stereotactic body radiation therapy,[19] from 61.5% to 44.8%.

In conclusion, we report QOL results from a prospective multicenter study of high-risk operable patients treated with SR. Global QOL measured with the SF36, and dyspnea measured with the SOBQ, did not deteriorate significantly after SR. Low baseline SF36 scores did not predict poor survival; however, low SOBQ scores at baseline, as well as a significant decline in SOBQ score at 12 months, did predict subsequent poor overall survival. Some advantages related to minimizing postoperative dyspnea, as measured by the SOBQ, were gained by using VATS (rather than thoracotomy) or wedge resection (rather

GTS

TABLE 3. Results of a landmark analysis for overall survival and recurrence-free survival using month-3 and month-12 quality-of-life scores

Outcome	Month 3		Month 12	
	10-point decline	No 10-point decline*	10-point decline	No 10-point decline*
SOBQ				
Overall survival				
N	45	127	38	104
Median (mo) (95% CI)	47.7 (30.4, NA)	60.2 (53.7, NA)	43.0 (16.3, NA)	NA (NA, NA)
HR (95% CI)	1.57 (0.93, 2.65)		2.10 (1.16, 3.81)	
P value	.09		.01	
3-y rate (%, 95% CI)†	65.2 (52.3, 81.1)	78.3 (71.3, 86.1)	63.9 (49.9, 81.8)	87.2 (81.0, 94.0)
Recurrence-free survival				
N	44	125	31	98
Median (mo) (95% CI)	47.7 (34.1, NA)	50.7 (38.1, NA)	43.0 (23.4, NA)	49.6 (41.7, NA)
HR (95% CI)	1.03 (0.63, 1.68)		1.39 (0.74, 2.61)	
P value	.90		.30	
2-y rate (%, 95% CI)†	75.1 (63.4, 89.0)	74.2 (66.9, 82.3)	62.7 (49.0, 80.3)	86.5 (80.2, 93.4)
SF36 physical component, gender and age adjusted				
Overall survival				
N	17	148	22	119
Median (mo) (95% CI)	NA (NA, NA)	60.2 (50.6, NA)	NA (NA, NA)	NA (NA, NA)
HR (95% CI)	1.06 (0.46, 2.47)		1.51 (0.73, 3.13)	
P value	.89		.26	
3-y rate (%, 95% CI)†	68.2 (48.6, 95.7)	75.7 (68.9, 83.2)	62.2 (44.6, 86.8)	85.3 (79.0, 92.0)
Recurrence-free survival				
N	15	147	15	113
Median (mo) (95% CI)	NA (NA, NA)	50.6 (37.9, 58.6)	34.6 (10.5, NA)	49.6 (42.4, NA)
HR (95% CI)	0.71 (0.29, 1.77)		1.52 (0.68, 3.39)	
P value	.47		.30	
2-y rate (%, 95% CI)†	70.1 (51.2, 96.0)	75.8 (69.1, 83.1)	54.2 (36.8, 79.8)	86.5 (80.6, 92.9)
SF36 mental component, gender and age adjusted				
Overall survival				
N	27	138	20	121
Median (mo) (95% CI)	60.2 (21.6, 60.2)	NA (NA, NA)	NA (NA, NA)	NA (NA, NA)
HR (95% CI)	1.40 (0.74, 2.65)		1.37 (0.62, 3.07)	
P value	.30		.44	
3-y rate (%, 95% CI)†	62.0 (45.0, 85.5)	77.3 (70.4, 84.8)	72.9 (54.9, 96.6)	83.2 (76.7, 90.2)
Recurrence-free survival				
N	27	135	17	111
Median (mo) (95% CI)	45.4 (14.8, NA)	50.7 (38.6, NA)	NA (NA, NA)	49.6 (41.7, NA)
HR (95% CI)	1.38 (0.79, 2.42)		0.80 (0.32, 2.01)	
P value	.26		.63	
2-y rate (%, 95% CI)†	65.2 (49.1, 86.4)	77.1 (70.3, 84.5)	74.7 (57.7, 96.6)	82.6 (76.2, 89.7)

P values are from the Cox Model. *SOBQ*, (University of California, San Diego) shortness of breath questionnaire; *NA*, not applicable; *CI*, confidence interval; *HR*, hazard ratio; *SF36*, 36-item short-form health survey. *Reference group. †Kaplan-Meier estimates using all data.

than segmentectomy). In addition, VATS, as opposed to thoracotomy, patients had improved PCS scores at 3 months, lending support to the preferential use of VATS when SR is performed.

Conflict of Interest Statement

Hiran C. Fernando reports consulting fees from Galil and CSA Medical. Bryan F. Meyers reports consulting fees from Ethicon and Varian. Joe B. Putnam reports consulting fees from GlaxoSmithKline. All other authors have nothing to disclose with regard to commercial support.

You can watch a Webcast of this AATS meeting presentation by going to: http://webcast.aats.org/2014/files/Tuesday/20140429_1120AM_1140AM_Hiran_Fernando_BryanF-Meyers.mp4

The authors thank the American College of Surgeons Oncology Group and Alliance staff, in particular the leadership of Heidi Nelson and David Ota for assistance in the completion of this project. The authors thank all of the investigators and their site research teams who enrolled patients in this study. These were at the following locations: University of Pittsburgh (Pittsburgh, Pa); Mayo Clinic (Rochester, Minn); Washington University

(St Louis, Mo); University of Virginia (Charlottesville, Va); Benedictine Hospital (Kingston, NY); University of Cincinnati (Cincinnati, Ohio); Jameson Hospital (New Castle, Pa); University of Michigan (Ann Arbor, Mich); Latter Day Saints Hospital (Salt lake City, Utah); Memorial Medical Center (Springfield, Ill); Rhode Island Hospital (Providence, RI); Valley Hospital (Ridgewood, NJ); William Beaumont Hospital (Royal Oak, Mich); Northwestern University (Chicago, Ill); Medical City Dallas (Dallas, Tex); Allegheny Cancer Center Network (Pittsburgh, Pa); Boston Medical Center (Boston, Mass); City of Hope Medical Center (Duarte, Calif); Portland Veterans Administration Medical Center (Portland, Ore); University of Philadelphia (Philadelphia, Pa); Virginia Mason Medical Center (Seattle, Wash); Medical University of South Carolina (Charleston, SC); Memorial Hospital (Chattanooga, Tenn); South Nassau Community Hospital (Oceanside, NY); Southern Illinois University School of Medicine (Springfield, Ill); Swedish Hospital (Seattle, Wash), University of Tennessee (Knoxville, Tenn), Dartmouth Hitchcock Medical Center (Lebanon, NH); Emory University (Atlanta, Ga); Fox Chase Cancer Center (Philadelphia, Pa); Oregon Health Sciences University (Portland, Ore); Vanderbilt University Medical Center (Nashville, Tenn); Intermountain Medical Center (Murray, Utah); London Health Sciences Centre (London, Ontario, Canada); Methodist Hospital (Houston, Tex); Miami Valley Hospital (Dayton, Ohio); Monmouth Medical Center (Long Branch, NJ); Northshore University Health System (Evanston, Ill); Providence Medical Center (Portland, Ore); Roswell Park Cancer Institute (Buffalo, NY); and Thomas Jefferson University Hospital (Philadelphia, Pa).

Finally, the authors thank the brave patients with non–small cell lung cancer and their caregivers who participated in this study.

References

1. Timmerman R, Paulus R, Galvin J, Michalski J, Straube W, Bradley J, et al. Sterotactic body radiation therapy for inoperable early stage lung cancer. *JAMA*. 2010;303:1070-6.
2. Dupuy DE, DiPetrillo T, Gandhi S, Ready N, Ng T, Donat W, et al. Radiofrequency ablation followed by conventional radiotherapy for medically inoperable stage I non-small cell lung cancer. *Chest*. 2006;129:738-45.
3. Fernando HC, Landreneau RJ, Mandrekar SJ, Nichols FC, Hillman SL, Heron DE, et al. The impact of brachytherapy on local recurrence rates after sublobar resection: Results from ACOSOG Z4032 (Alliance), a phase III randomized trial for high-risk operable non-small cell lung cancer (NSCLC). *J Clin Oncol*. 2013;31(15 Suppl). Abst. 7502.
4. D'Amato TA, Galloway M, Szydlowski G, Chen A, Landreneau RJ. Intraoperative brachytherapy following thoracoscopic wedge resection of stage I lung cancer. *Chest*. 1998;114:1112-5.
5. Lee W, Daly BD, DiPetrillo TA, Morelli DM, Neuschatz AC, Morr J, et al. Limited resection for non-small cell lung cancer: observed local control with implantation of I-125 brachytherapy seeds. *Ann Thorac Surg*. 2003;75:237-43.
6. US Department of Health and Human Services (USDHHS), National Institutes of Health, National Cancer Institute. *Common terminology criteria for adverse events (CTCAE), version 3*. Bethesda, Md: USDHHS; 2003.
7. Fernando HC, Landreneau RJ, Mandrekar SJ, Hillman SL, Nichols FC, Meyers B, et al. Thirty and ninety-day outcomes after sublobar resection with and without brachytherapy for non-small cell lung cancer: results from a multicenter phase III study. *J Thorac Cardiovasc Surg*. 2011;142:1143-51.
8. Ware JE Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assesment (IQOLA) Project. *J Clin Epidemiol*. 1998;51:903-12.
9. Swigris JJ, Han M, Vij R, Noth I, Eisenstein EL, Anstrom KJ, et al. The UCSD shortness of breath questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. *Respir Med*. 2012;106:1447-55.
10. Thalji L, Haggerty CC, Rubin R, Berckmans TR, Pardee BL. *1990 National Survey of Functional Health Status, Final Report*. Chicago, Ill: National Opinion Research Center; 1991.
11. Sloan JA, Cella D, Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. *J Clin Epidemiol*. 2005;58:1217-9.
12. Handy JR Jr, Asaph JW, Douville EC, Ott GY, Grunkemeier GL, Wu Y. Does video assisted thoracoscopic lobectomy for lung cancer provide improved functional outcomes compared with open lobectomy? *Eur J Cardiothorac Surg*. 2010;37:451-5.
13. Brunelli A, Salati M, Refai M, Xiumé F, Berardi R, Mazzanti P, et al. Development of patient-centered aggregate score to predict survival after lung resection for non-small cell lung cancer. *J Thorac Cardiovasc Surg*. 2013;146:385-90.
14. Balduyck B, Hendriks J, Lauwers P, Van Schil P. Quality of life evolution after lung cancer surgery: a prospective study in 100 patients. *Lung Cancer*. 2007;56:423-31.
15. Kenny PM, King MT, Viney RC, Boyer MJ, Pollicino CA, McLean JM, et al. Quality of life and survival in the 2 years after surgery for non-small cell lung cancer. *J Clin Oncol*. 2008;26:233-41.
16. El-Sherif A, Fernando HC, Santos R, Pettiford B, Luketich JD, Close JM, et al. Margin and local recurrence after sublobar resection of non-small cell lung cancer. *Ann Surg Oncol*. 2007;14:2400-5.
17. Brunelli A, Socci L, Refai M, Salati M, Xiumé F, Sabbatini A. Quality of life before and after major lung resection for lung cancer: a prospective follow-up analysis. *Ann Thorac Surg*. 2007;84:410-6.
18. Brown SR, Mathew R, Keding A, Marshall HC, Brown JM, Jayne DG. The impact of post-operative complications on long-term quality of life after curative colorectal cancer surgery. *Ann Surg*. 2014;259:916-23.
19. Videtic GM, Chandana AR, Sorenson L. A postoperative study of quality of life including fatigue and pulmonary function after stereotactic body radiotherapy for medically inoperable early-stage lung cancer. *Support Care Cancer*. 2013;21:211-8.

Key words: lung cancer, surgery, quality of life

Discussion

Dr Michael Jaklitsch *(Boston, Mass).* We are all indebted to the authors for accumulating preoperative and serial postoperative quality-of-life scores in 212 patients in a multi-institutional prospective randomized trial of sublobar resection versus sublobar resection with brachytherapy in a high-risk surgical population.

Quality-of-life measures were obtained using the SF36 and SOBQ dyspnea scale. It is the only prospectively accumulated data that I know of that has been reported within a high-risk operable group for which the planned operation was always limited to wedge or segment, owing to the baseline functional status. It reports important benchmark measures that can be compared with other modalities of therapy and has great value in that regard.

However, the hypothesis presented in the article is that one of the preoperative scores would predict which patients were at highest risk for complications, and the SF36 had neither predictive nor prognostic value in this study. The SOBQ dyspnea score had both predictive and prognostic values, but not with a robust association.

The SF36 is relatively quick, but I have found it too crude to be discriminating in elderly patients. Time and again, we find the greatest value in the simplest preoperative measures, namely, 6-minute walk distance, stair climbing, and the ability to care for oneself in activities of daily living. Data from these 3 tests—6-minute walk distance, stair climbing, and activities of daily living—are not included in this analysis. Was there an intention at trial design to compare the SF36 and the SOBQ data with these tests?

Secondly, preoperative performance status was measured, judging by Table 1, yet I did not see an analysis of the ability of performance status measures to predict adverse outcomes. You have a large proportion of impaired performance status patients within the trial, so that analysis could be performed.

Finally, I would like to make just a comment to ask why it is a negative trial. There are plenty of adverse events in the trial, and it is an impaired case. Is your theory, the mathematical theory, that you can predict—within this universe of high-probability—adverse outcomes with these measures? Then, the question becomes, going back to the earlier question—if you were to give advice to thoracic surgeons here at this meeting, would you have them add the SF36 and the SOBQ dyspnea score to what they already do, or would you advise them to continue to depend on the 6-minute walk distance, the stair climbing, and activities of daily living?

Thank you very much for the privilege of reviewing the manuscript.

Dr Meyers. Thanks, Dr Jaklitsch. You have a point that, with regard to clinical care of patients, it is quite reasonable to limit yourself to things like activities of daily living or a 6-minute walk distance—something that could be very easily measurable in clinical practice.

With regard to performing a clinical trial, when we are creating a clinical trial, there are always voices throughout the room trying to add 1 more test and 1 more survey until the clinical trial becomes so onerous for those in the trial and for the patient that it becomes a challenge to conduct. It is always an exercise of discretion to pick instruments to measure quality of life, but we always have to leave some behind.

I do not believe that the primary plan of this trial was to use quality-of-life scores to predict adverse outcomes. Prediction is always kind of a risky claim. There was an interest in looking for an association between different levels of quality of life and adverse outcomes; but the quality-of-life endpoint was originally included because we knew that the survival outcome of the brachytherapy intervention was likely to be close to that of the nonbrachytherapy intervention. If you have 2 interventions that have a very similar or identical outcome with regard to survival or recurrence, then you might pick a winner by another dimension, which in this case would be quality of life.

As it turned out, the freedom from recurrence was identical, the survival was identical, and the quality of life was identical. So it turned out not to be a tiebreaker or a discriminator between the 2 therapies.

Dr Todd L. Demmy *(Buffalo, NY)*. A common thread in quality-of-life research is that with less lung after resection, people tend to experience more pain. Were you able to tease out the component of pain in your quality-of-life assessments to see if that is perhaps why the VATS had a bigger effect in that group?

Thank you.

Dr Meyers. As you know, the SF36 has a physical component score and a mental component score. I imagine pain impacts both of them; it affects mobility and general outlook. We did not have a specific separate pain score that we were assessing.