A Randomized Trial Comparing Lung-Volume–Reduction Surgery with Medical Therapy for Severe Emphysema

National Emphysema Treatment Trial Research Group*

ABSTRACT

BACKGROUND

Lung-volume–reduction surgery has been proposed as a palliative treatment for severe emphysema. Effects on mortality, the magnitude and durability of benefits, and criteria for the selection of patients have not been established.

METHODS

A total of 1218 patients with severe emphysema underwent pulmonary rehabilitation and were randomly assigned to undergo lung-volume–reduction surgery or to receive continued medical treatment.

RESULTS

Overall mortality was 0.11 death per person-year in both treatment groups (risk ratio for death in the surgery group, 1.01; P=0.90). After 24 months, exercise capacity had improved by more than 10 W in 15 percent of the patients in the surgery group, as compared with 3 percent of patients in the medical-therapy group (P<0.001). With the exclusion of a subgroup of 140 patients at high risk for death from surgery according to an interim analysis, overall mortality in the surgery group was 0.09 death per person-year, as compared with 0.10 death per person-year in the medical-therapy group (risk ratio, 0.89; P=0.31); exercise capacity after 24 months had improved by more than 10 W in 16 percent of patients in the surgery group, as compared with 3 percent of patients in the medical-therapy group (P<0.001). Among patients with predominantly upper-lobe emphysema and low exercise capacity, mortality was lower in the surgery group than in the medical-therapy group (risk ratio for death, 0.47; P=0.005). Among patients with non–upper-lobe emphysema and high exercise capacity, mortality was higher in the surgery group than in the medical-therapy group (risk ratio, 2.06; P=0.02).

CONCLUSIONS

Overall, lung-volume–reduction surgery increases the chance of improved exercise capacity but does not confer a survival advantage over medical therapy. It does yield a survival advantage for patients with both predominantly upper-lobe emphysema and low base-line exercise capacity. Patients previously reported to be at high risk and those with non–upper-lobe emphysema and high base-line exercise capacity are poor candidates for lung-volume–reduction surgery, because of increased mortality and negligible functional gain.

The writing committee for the National Emphysema Treatment Trial (NETT) (Alfred Fishman, M.D., University of Pennsylvania, Philadelphia; Fernando Martinez, M.D., University of Michigan, Ann Arbor; Keith Naunheim, M.D., Saint Louis University, St. Louis; Steven Piantadosi, M.D., Ph.D., and Robert Wise, M.D., Johns Hopkins University, Baltimore; Andrew Ries, M.D., M.P.H., University of California, San Diego, La Jolla; Gail Weinmann, M.D., National Heart, Lung, and Blood Institute, Bethesda, Md.; and Douglas E. Wood, M.D., University of Washington, Seattle) takes responsibility for the content of this article. Address reprint requests to Dr. Piantadosi at the NETT Coordinating Center, 615 N. Wolfe St., Rm. 5010, Baltimore, MD 21205.

*The members of the National Emphysema Treatment Trial Research Group are listed in the Appendix.

LUNG-VOLUME–REDUCTION SURGERY has been proposed as a palliative treatment for patients with severe emphysema.[1-8] Uncertainty about morbidity and mortality; the occurrence, magnitude, and duration of benefit; and preoperative predictors of benefit led us to conduct a federally sponsored, multicenter, randomized clinical trial, the National Emphysema Treatment Trial (NETT).[9] The primary outcomes for the trial were mortality and maximal exercise capacity two years after randomization. Secondary outcomes included the distance walked in six minutes, pulmonary function, quality of life, and degree of dyspnea. An important goal of the trial was to identify patient-selection criteria for lung-volume–reduction surgery. Criteria for inclusion were crafted to include all patients who might benefit from lung-volume–reduction surgery. Here, we report outcomes for all patients who underwent randomization and identify subgroups of patients with varying levels of risk and benefit.

METHODS

The design and methods of the trial have been described previously.[9] We summarize them below.

POPULATION OF PATIENTS AND ASSESSMENTS

At 17 clinics, patients with severe emphysema underwent comprehensive medical evaluation to ensure compliance with usual medical therapy and to rule out clinically significant coexisting conditions.[9] Base-line measurements were completed after pulmonary rehabilitation but before randomization, and patients underwent complete evaluations at 6 months, 12 months, and yearly thereafter. Overall mortality and maximal exercise capacity (on cycle ergometry with an increment of 5 or 10 W per minute after three minutes of pedaling with the ergometer set at 0 W and the patient breathing 30 percent oxygen) were the primary outcome measures. Secondary outcome measures included pulmonary function,[10,11] the distance walked in six minutes,[12,13] and the results on a self-administered questionnaire about health-related quality of life (St. George's Respiratory Questionnaire),[14] a general quality-of-life questionnaire (the Quality of Well-Being scale),[15] and a dyspnea questionnaire (the University of California, San Diego, Shortness of Breath Questionnaire[16]).

The distribution of emphysema was classified as heterogeneous or homogeneous on the basis of high-resolution computed tomography (CT) with the use of a visual scoring system.[17] In addition, the radiologist classified the craniocaudal distribution of emphysema as predominantly affecting the upper lobes, predominantly affecting the lower lobes, diffuse, or predominantly affecting the superior segments of the lower lobes; the latter three categories were grouped together as predominantly non–upper-lobe emphysema for the purposes of our analysis. The ratio of perfusion in the upper regions of the lungs to that in the lower regions was quantified on the basis of radionuclide scans of the lungs that were interpreted at each center.

Before randomization, eligible patients completed 6 to 10 weeks of pulmonary rehabilitation supervised by study personnel. All patients provided written informed consent, and the study was approved by the institutional review board at each clinic. In May 2001, patients with a forced expiratory volume in one second (FEV_1) that was 20 percent or less of the predicted value and either a homogeneous distribution of emphysema or a carbon monoxide diffusing capacity that was 20 percent or less of the predicted value were determined to be at high risk for death after lung-volume–reduction surgery, with a low probability of functional benefit,[17] and were no longer eligible for randomization. Detailed criteria for inclusion in and exclusion from the trial are listed in Supplementary Appendix 1 (available with the full text of this article at http://www.nejm.org).

Patients randomly assigned to lung-volume–reduction surgery underwent bilateral stapled wedge resection through a median sternotomy or video-assisted thoracic surgery; the goal was to resect 20 to 35 percent of each lung, targeting the most diseased portions. Eight centers performed lung-volume reduction by median sternotomy alone, three centers by video-assisted thoracic surgery alone, and six by median sternotomy or video-assisted thoracic surgery selected randomly. Patients' adherence to medical regimens, abstinence from tobacco use, and pulmonary-rehabilitation treatment at home were monitored through regular telephone calls and clinic visits.

DEFINITION OF OUTCOMES

Vital status, ascertained as of December 2002, was determined by reports from the clinical centers and review of the Social Security Administration's December 2002 Death Master File.[18,19] Total, 30-day, and 90-day mortality from all causes were measured

from the day of randomization for both treatment groups. On the basis of previous experience, we defined improvement in maximal exercise capacity as an increase in the maximal workload of more than 10 W from the postrehabilitation base-line level. Improvement in health-related quality of life was defined as a decrease in the score on the St. George's Respiratory Questionnaire of more than 8 points (on a 100-point scale) from the base-line level (measured after rehabilitation). These thresholds are higher than those typically used to define minimal clinically important differences (e.g., a four-point change on the respiratory questionnaire is usually used[20]) but were selected to represent a degree of improvement that would be appropriate to justify the high risks associated with surgery in patients with severe emphysema. Patients who died or were missing data required for the assessment were considered not to have had improvement.

STATISTICAL ANALYSIS

All analyses were performed according to the intention-to-treat principle. Fisher's exact test was used to compare the proportions of patients in each group who died.[21] The trial protocol specified that the primary comparison would be between the proportions of patients in the two groups who died rather than a comparison with the use of a typical rank-based test (e.g., a log-rank or Wilcoxon test) for differences in mortality, because the hazard functions for death were expected to cross, resulting in nonproportional hazards. This crossing of hazards was anticipated as a consequence of the expected perioperative mortality, potentially followed by lower mortality in the surgery group after six months.[22] The risk ratio for death was estimated on the basis of the overall mortality in each group after a mean of 29.2 months of follow-up.[21]

Despite randomization, differential early mortality makes survivors in the surgery group appear healthier than their counterparts in the medical-therapy group. This imbalance confounds the interpretation of outcomes measured only in survivors, such as exercise capacity, pulmonary function, and quality of life. Therefore, we used classifications derived from measured outcomes defined for all patients, such as improvement versus no improvement, with the latter category also including patients who were unable to complete the evaluation or who had died. The cutoff point used for a derived variable — for example, the definition of improvement — is necessarily somewhat arbitrary.

Histograms were used to compare the surgery group and the medical-therapy group in terms of the distributions of categories of change from base line in the outcome measures; all patients who were followed for 6, 12, and 24 months were included in these analyses. The histograms included 8 or 10 categories of the change measured in survivors (scored from 2 to 12, with higher scores indicating more improvement), a category for patients who died (scored as 0), and a category for patients who missed or were unable to complete the evaluation (scored as 1). For cycle ergometry, patients who could pedal only with the ergometer set at 0 W were also given a score of 1. For the Quality of Well-Being scale, patients who died were assigned a score of 0 on the questionnaire for that visit, and patients who did not complete the questionnaire were assigned a score equal to half of the lowest score observed for that visit. P values for the comparisons of the two groups in terms of the distribution of scores were derived by the Wilcoxon rank-sum test.[23]

Subgroups of patients with differential risks or benefits were identified with the use of a series of logistic-regression analyses that included as the outcome death, improvement in exercise capacity, or improvement in health-related quality of life at 6, 12, or 24 months. For each outcome, the model included terms for the treatment-group assignment, a base-line prognostic factor, and the interaction between the treatment group and the prognostic factor, to test whether or not the factor was associated with differential outcome. P values for the interaction terms were determined by exact-score tests for logistic regression,[24] and factors with statistically significant interactions were also examined in logistic models including pairwise combinations of the factors with corresponding interaction terms.

Most prognostic factors were identified by a hypothesis specified in the trial protocol (age, percentage of the predicted value for the FEV_1, partial pressure of arterial carbon dioxide, percentage of the predicted value for the residual volume, distribution of perfusion on radionuclide scanning of the lungs, homogeneity or heterogeneity of the distribution of emphysema on CT, and presence or absence of hyperinflation on chest radiography).[17] Other base-line prognostic factors were added by the data and safety monitoring board or the investigators after the initiation of the trial but well before the completion of data collection (percentage of

predicted value for carbon monoxide diffusing capacity, maximal exercise capacity, ratio of residual volume to total lung capacity, ratio of expired ventilation in one minute to carbon dioxide excretion in one minute, presence or absence of upper-lobe predominance of emphysema, degree of dyspnea, quality of life, race or ethnic group, and sex). Measured prognostic factors were categorized into approximate quartiles, and patients in the quartile with the worst prognosis were compared with all other patients. The prognostic factors were analyzed quarterly as part of ongoing monitoring of the study to meet the a priori objective of identifying characteristics of patients who might have a differential risk of harm or differential benefit from lung-volume–reduction surgery.

A base-line maximal exercise capacity in the lowest quartile ($\leq$25 W) was one of the two prognostic factors that identified patients with a differential risk of death. Recognizing that there are differences between men and women in exercise capacity, we refined the cutoff point for base-line maximal exercise capacity by examining a range of sex-specific cutoff points, and we found that the sex-specific 40th percentile (25 W for women and 40 W for men) was the best cutoff point for the classification of patients with a differential risk of death from lung-volume–reduction surgery (see Supplementary Appendix 2, available with the full text of this article at http://www.nejm.org).

All reported P values are based on two-sided tests. P values were not corrected for multiple comparisons. The primary and secondary objectives of the trial were identified by hypothesis before the study began and specifically included the identification of subgroups associated with differential harm or benefit from lung-volume–reduction surgery. Although a large number of statistical tests were possible, we limited the number through planning and an orderly analysis of the data. The prognostic factors identified have a plausible clinical rationale and were associated with large, statistically significant risk ratios.

The study protocol specified that recruitment would end by July 2002, with an accrual of 2500 patients and study completion in December 2002. This design was based on a primary survival comparison allowing for 8 percent mortality in the group assigned to medical therapy and for a rate of unplanned crossover of 30 percent among those assigned to medical therapy.[9] The trial was ended on the planned date despite accrual that was lower than expected. Had the study been designed with the assumptions of the higher mortality rate and the lower crossover rate that were actually observed (0.11 death per person-year and 5.4 percent, respectively), the recruitment goal would have been 1190 patients.

RESULTS

STUDY PATIENTS

Between January 1998 and July 2002, 3777 patients were evaluated, and 1218 patients underwent randomization — 608 to surgery and 610 to medical therapy. The groups had similar base-line characteristics after rehabilitation, except that there was a higher proportion of men in the medical-therapy group (Table 1) (Results for subgroups are given in Supplementary Appendix 3 [available with the full text of this article at http://www.nejm.org]). As of December 2002, 99 percent of surviving patients continued to complete quarterly telephone interviews or annual clinic visits.

TREATMENT

Of the 608 patients assigned to lung-volume–reduction surgery, 580 (95.4 percent) underwent surgery (406 [70.0 percent] of them by median sternotomy and 174 [30.0 percent] by video-assisted thoracic surgery), 21 (3.5 percent) declined to undergo surgery, and 7 (1.2 percent) were deemed to be unsuitable for surgery after randomization. The median time from randomization to surgery was 10 days; 74 of the patients assigned to lung-volume–reduction surgery (12.2 percent) underwent surgery more than 14 days after randomization. Deviations from the surgical protocol (unilateral surgery or bilateral surgery performed in two sessions) occurred in 12 patients (2.0 percent) because of intraoperative factors; 4 patients in the surgery group (0.7 percent) received lung transplants after undergoing lung-volume–reduction surgery. Among the 610 patients assigned to medical therapy, 33 (5.4 percent) underwent lung-volume–reduction surgery outside the study, and 15 (2.5 percent) received lung transplants during follow-up.

OUTCOMES FOR ALL 1218 PATIENTS

The 90-day mortality rate in the surgery group was 7.9 percent (95 percent confidence interval, 5.9 to 10.3) and was significantly higher than that in the medical-therapy group (1.3 percent [95 percent confidence interval, 0.6 to 2.6], P<0.001) (Table 2).

Table 1. Characteristics of All 1218 Patients at Base Line.*		
Characteristic	Surgery Group (N=608)	Medical-Therapy Group (N=610)
Age at randomization — yr	66.5±6.3	66.7±5.9
Race or ethnic group — no. (%)		
Non-Hispanic white	581 (96)	575 (94)
Non-Hispanic black	19 (3)	23 (4)
Other	8 (1)	12 (2)
Sex — no. (%)†		
Female	253 (42)	219 (36)
Male	355 (58)	391 (64)
Distribution of emphysema on CT — no. (%)‡		
Predominantly upper lobe	385 (63)	405 (67)
Predominantly non–upper lobe	223 (37)	204 (33)
Heterogeneous	330 (54)	336 (55)
Homogeneous	278 (46)	274 (45)
Perfusion ratio§	0.30±0.21	0.28±0.23
Maximal workload — W	38.7±21.1	39.4±22.2
Distance walked in 6 min — ft¶	1216.5±312.6	1219.0±316.0
FEV_1 after bronchodilator use — % of predicted value	26.8±7.4	26.7±7.0
Total lung capacity after bronchodilator use — % of predicted value	128.0±15.3	128.5±15.0
Residual volume after bronchodilator use — % of predicted value	220.5±49.9	223.4±48.9
Carbon monoxide diffusing capacity — % of predicted value	28.3±9.7	28.4±9.7
PaO_2 — mm Hg	64.5±10.5	64.2±10.1
$PaCO_2$ — mm Hg	43.3±5.9	43.0±5.8
Total score on St. George's Respiratory Questionnaire‖	52.5±12.6	53.6±12.7
Average daily Quality of Well-Being score**	0.58±0.12	0.56±0.11
Total UCSD Shortness of Breath score††	61.6±18.1	63.4±18.6

* Base-line measurements were obtained after rehabilitation but before randomization, except for the carbon monoxide diffusing capacity, which was measured before rehabilitation. Plus–minus values are means ±SD. CT denotes computed tomography, FEV_1 forced expiratory volume in one second, PaO_2 partial pressure of arterial oxygen, and $PaCO_2$ partial pressure of arterial carbon dioxide.
† P for homogeneity=0.04.
‡ Upper-lobe predominance of emphysema was judged subjectively by each center's radiologist, who described the distribution of disease as predominantly upper lobe, predominantly lower lobe, diffuse, or predominantly affecting superior segments of the lower lobes. The latter three choices were grouped as predominantly non–upper lobe. The classification of the emphysema as heterogeneous or homogeneous was based on subjective scores assigned by each center's radiologist to each of three zones in each lung. Data on upper-lobe versus non–upper-lobe distribution were missing for one patient.
§ The perfusion ratio is derived from the radionuclide perfusion scan. Each lung is divided into three zones, and a percentage of total perfusion is assigned to each zone. The ratio is calculated as the sum of the percentages assigned to the two upper zones divided by the sum of the per-centages assigned to the four middle and lower zones.
¶ To convert values from feet to meters, divide by 3.28.
‖ The St. George's Respiratory Questionnaire is a 51-item questionnaire on the health-related quality of life with regard to respiratory symp-toms that is completed by the patient; the total score ranges from 0 to 100, with lower scores indicating better health-related quality of life.
** The Quality of Well-Being scale is a 77-item quality-of-life questionnaire completed by the patient. The average daily score ranges from 0 to 1, with higher scores indicating better quality of life.
†† The University of California, San Diego (UCSD), Shortness of Breath Questionnaire is a 24-item questionnaire about dyspnea that is com-pleted by the patient; the total score ranges from 0 to 120, with lower scores indicating less shortness of breath.

The 90-day mortality rate was similar among patients who underwent video-assisted thoracic surgery and among those who underwent median sternotomy (6.1 percent and 8.6 percent, respectively; P=0.33). All clinics reported similar rates of mortality, morbidity, and common intraoperative and postoperative complications. During follow-up (mean duration, 29.2 months), 157 patients assigned to lung-volume–reduction surgery and 160 assigned to medical therapy died. The total mortality rate was 0.11 death per person-year in both groups (risk ratio for death in the surgery group, 1.01; P=0.90) (Table 2). There was no significant difference in overall mortality despite a higher early mortality rate in the surgery group (Fig. 1A).

Exercise capacity improved by more than 10 W in 28 percent, 22 percent, and 15 percent of patients in the surgery group after 6, 12, and 24 months, respectively, as compared with 4 percent, 5 percent, and 3 percent of patients in the medical-therapy group (P<0.001 for the comparisons at all three time points) (Table 3 and Supplementary Appendix 4 [available with the full text of this article at http://www.nejm.org]). Patients in the surgery group were significantly more likely to have improvements than patients in the medical-therapy group in the distance walked in six minutes, percentage of the predicted value for FEV_1, general and health-related quality of life, and degree of dyspnea (see Supplementary Appendix 4).

HIGH-RISK PATIENTS

The subgroup of 140 patients with a value for FEV_1 that was 20 percent or less of the predicted value

Table 2. Mortality among All Patients and in Subgroups.*

Patients	90-Day Mortality			Total Mortality					
	Surgery Group	Medical-Therapy Group	P Value	Surgery Group		Medical-Therapy Group		Risk Ratio	P Value
				no. of deaths/ total no.	no. of deaths/ person-yr	no. of deaths/ total no.	no. of deaths/ person-yr		
	no. of deaths/total no. (% [95% CI])								
All patients	48/608 (7.9 [5.9–10.3])	8/610 (1.3 [0.6–2.6])	<0.001	157/608	0.11	160/610	0.11	1.01	0.90
High-risk†	20/70 (28.6 [18.4–40.6])	0/70 (0 [0–5.1])	<0.001	42/70	0.33	30/70	0.18	1.82	0.06
Other	28/538 (5.2 [3.5–7.4])	8/540 (1.5 [0.6–2.9])	0.001	115/538	0.09	130/540	0.10	0.89	0.31
Subgroups‡									
Patients with predominantly upper-lobe emphysema									
Low exercise capacity	4/139 (2.9 [0.8–7.2])	5/151 (3.3 [1.1–7.6])	1.00	26/139	0.07	51/151	0.15	0.47	0.005
High exercise capacity	6/206 (2.9 [1.1–6.2])	2/213 (0.9 [0.1–3.4])	0.17	34/206	0.07	39/213	0.07	0.98	0.70
Patients with predominantly non–upper-lobe emphysema									
Low exercise capacity	7/84 (8.3 [3.4–16.4])	0/65 (0 [0–5.5])	0.02	28/84	0.15	26/65	0.18	0.81	0.49
High exercise capacity	11/109 (10.1 [5.1–17.3])	1/111 (0.9 [0.02–4.9])	0.003	27/109	0.10	14/111	0.05	2.06	0.02

* Mortality was measured from the date of randomization in both treatment groups. Total mortality rates are based on a mean follow-up of 29.2 months. P values were calculated by Fisher's exact test. Risk ratios are for the risk in the surgery group as compared with the risk in the medical-therapy group. A low base-line exercise capacity was defined as a postrehabilitation base-line maximal workload at or below the sex-specific 40th percentile (25 W for women and 40 W for men); a high-exercise capacity was defined as a workload above this threshold. CI denotes confidence interval.

† High-risk patients were defined as those with a forced expiratory volume in one second (FEV_1) that was 20 percent or less of the predicted value and either homogeneous emphysema on computed tomography or a carbon monoxide diffusing capacity that was 20 percent or less of the predicted value.

‡ High-risk patients were excluded from the subgroup analyses. For total mortality, P for interaction=0.004; this P value was derived from binary logistic-regression models with terms for treatment, subgroup, and the interaction between the two, with the use of an exact-score test with three degrees of freedom. Other factors that were considered as potential variables for the definition of subgroups included the base-line FEV_1, carbon monoxide diffusing capacity, partial pressure of arterial carbon dioxide, residual volume, ratio of residual volume to total lung capacity, ratio of expired ventilation in one minute to carbon dioxide excretion in one minute, distribution of emphysema (heterogeneous vs. homogeneous), perfusion ratio, score for health-related quality of life, and Quality of Well-Being score; age; race or ethnic group; and sex.
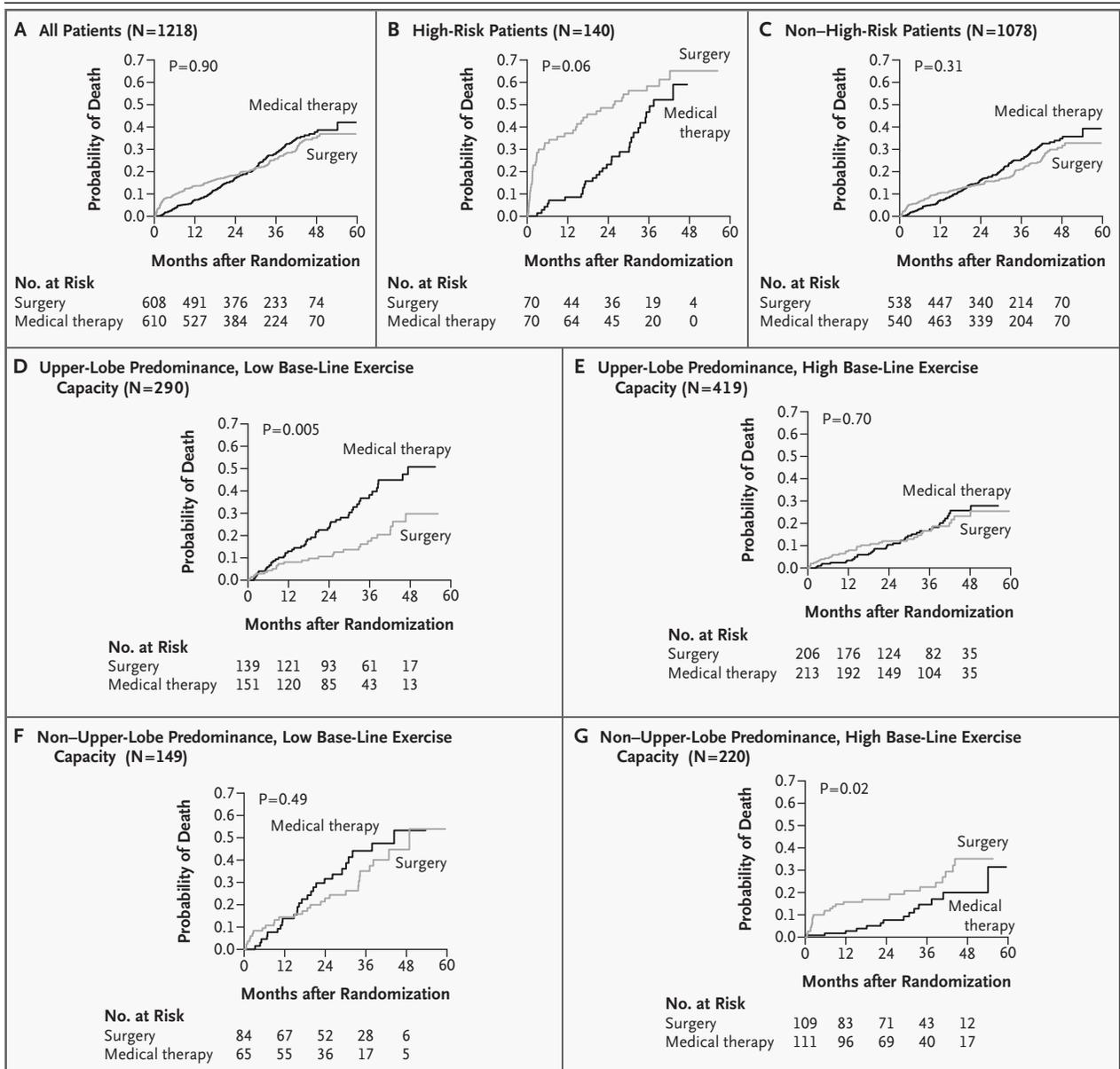
Figure 1. Kaplan–Meier Estimates of the Probability of Death as a Function of the Number of Months after Randomization.

P values were derived by Fisher's exact test for the comparison between groups over a mean follow-up period of 29.2 months. High-risk patients were defined as those with a forced expiratory volume in one second that was 20 percent or less of the predicted value and either homogeneous emphysema or a carbon monoxide diffusing capacity that was 20 percent or less of the predicted value. A low base-line exercise capacity was defined as a maximal workload at or below the sex-specific 40th percentile (25 W for women and 40 W for men); a high exercise capacity was defined as a workload above this threshold. This was an intention-to-treat analysis.

and either homogeneous emphysema or a carbon monoxide diffusing capacity that was 20 percent or less of the predicted value was previously reported to be at high risk for death after lung-volume–reduction surgery, with little chance of functional benefit.[17] The updated analyses of mortality and func-

tional improvement in this subgroup support the previous findings (Tables 2 and 3 and Fig. 1B).

OUTCOMES FOR PATIENTS WITHOUT HIGH RISK

Among the 1078 patients who were not at high risk, the 30-day mortality rate was 2.2 percent in

Table 3. Improvement in Exercise Capacity and Health-Related Quality of Life at 24 Months.*

Patients	Improvement in Exercise Capacity				Improvement in Health-Related Quality of Life			
	Surgery Group	Medical-Therapy Group	Odds Ratio	P Value	Surgery Group	Medical-Therapy Group	Odds Ratio	P Value
	no./total no. (%)				*no./total no. (%)*			
All patients	54/371 (15)	10/378 (3)	6.27	<0.001	121/371 (33)	34/378 (9)	4.90	<0.001
High-risk†	4/58 (7)	1/48 (2)	3.48	0.37	6/58 (10)	0/48	—	0.03
Other	50/313 (16)	9/330 (3)	6.78	<0.001	115/313 (37)	34/330 (10)	5.06	<0.001
Subgroups‡								
Predominantly upper-lobe emphysema								
Low exercise capacity	25/84 (30)	0/92	—	<0.001	40/84 (48)	9/92 (10)	8.38	<0.001
High exercise capacity	17/115 (15)	4/138 (3)	5.81	0.001	47/115 (41)	15/138 (11)	5.67	<0.001
Predominantly non–upper-lobe emphysema								
Low exercise capacity	6/49 (12)	3/41 (7)	1.77	0.50	18/49 (37)	3/41 (7)	7.35	0.001
High exercise capacity	2/65 (3)	2/59 (3)	0.90	1.00	10/65 (15)	7/59 (12)	1.35	0.61

* Improvement in exercise capacity in patients followed for 24 months after randomization was defined as an increase in the maximal workload of more than 10 W from the patient's postrehabilitation base-line value. Improvement in the health-related quality of life in patients followed for 24 months after randomization was defined as a decrease in the score on the St. George's Respiratory Questionnaire of more than 8 points (on a 100-point scale) from the patient's postrehabilitation base-line score. For both analyses, patients who died or who missed the 24-month assessment were considered not to have improvement. Odds ratios are for improvement in the surgery group as compared with the medical-therapy group. P values were calculated by Fisher's exact test. A low base-line exercise capacity was defined as a postrehabilitation base-line maximal workload at or below the sex-specific 40th percentile (25 W for women and 40 W for men); a high exercise capacity was defined as a workload above this threshold.

† High-risk patients were defined as those with a forced expiratory volume in one second (FEV_1) that was 20 percent or less of the predicted value and either homogeneous emphysema on computed tomography or a carbon monoxide diffusing capacity that was 20 percent or less of the predicted value.

‡ High-risk patients were excluded from the subgroup analyses. For improvement in exercise capacity, P for interaction=0.005; for improvement in health-related quality of life, P for interaction=0.03. These P values were derived from binary logistic-regression models with terms for treatment, subgroup, and the interaction between the two, with the use of an exact-score test with three degrees of freedom. Other factors that were considered as potential variables for the definition of subgroups included the base-line FEV_1, carbon monoxide diffusing capacity, partial pressure of arterial carbon dioxide, residual volume, ratio of residual volume to total lung capacity, ratio of expired ventilation in one minute to carbon dioxide excretion in one minute, distribution of emphysema (heterogeneous vs. homogeneous), perfusion ratio, score for health-related quality of life, and Quality of Well-Being score; age; race or ethnic group; and sex.

the surgery group, as compared with 0.2 percent in the medical-therapy group (P<0.001), and the 90-day mortality rate was 5.2 percent in the surgery group, as compared with 1.5 percent in the medical-therapy group (P=0.001) (Table 2). One month after randomization, 28.1 percent of the patients in the surgery group, as compared with 2.2 percent of the patients in the medical-therapy group, were hospitalized, living in a nursing or rehabilitation facility, or unavailable for interview but not known to be dead (P<0.001 for the comparison between groups); at two months, the percentages were 14.3 percent and 3.3 percent, respectively (P<0.001); at four months, 6.7 percent and 3.2 percent, respectively (P=0.007); and at eight months, 3.3 percent and 3.7 percent, respectively (P=0.87) (see Supple-

mentary Appendix 5 [available with the full text of this article at http://www.nejm.org]).

Total mortality among patients who were not at high risk during the trial was 0.09 death per person-year in the surgery group, as compared with 0.10 death per person-year in the medical-therapy group (risk ratio, 0.89; P=0.31) (Table 2 and Fig. 1C). Changes in exercise capacity, distance walked in six minutes, percentage of the predicted value for FEV_1, quality of life, and degree of dyspnea at 6, 12, and 24 months all favored the surgery group (Table 3, Fig. 2, and Supplementary Appendix 6 [available with the full text of this article at http://www.nejm.org]). When the analysis was limited to survivors who were able to complete the follow-up assessments, the pattern of changes in the outcome
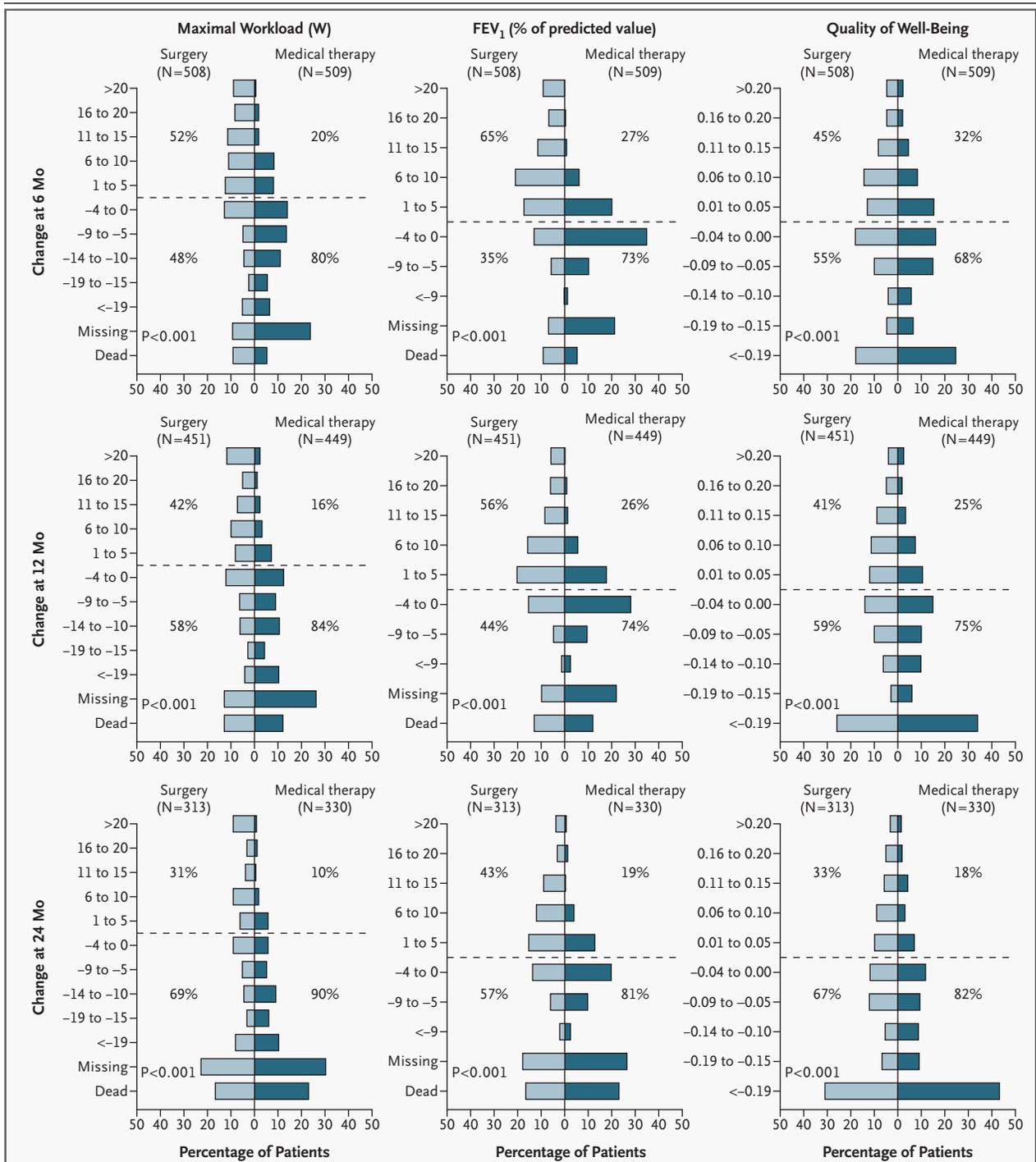
Figure 2. Histograms of Changes from Base Line in Exercise Capacity (Maximal Workload), Percentage of the Predicted Value for Forced Expiratory Volume in One Second (FEV$_1$), and Quality of Life (Quality of Well-Being Score) after 6, 12, and 24 Months of Follow-up.

Base-line measurements were performed after pulmonary rehabilitation. Patients previously identified as high-risk were excluded. Patients who were too ill to complete the procedure or who declined to complete the procedure but did not explain why were included in the "missing" category. For the Quality of Well-Being score, patients who died were assigned a score of 0 on the questionnaire for the visit, and patients who did not complete the questionnaire were assigned a score equal to half of the lowest score observed for the visit. P values were determined by the Wilcoxon rank-sum test. The degree to which the bars are shifted to the upper left of the chart indicates the degree of relative benefit of lung-volume–reduction surgery over medical treatment. The percentage shown in each quadrant is the percentage of patients in the specified treatment group with a change in the outcome falling into that quadrant. This was an intention-to-treat analysis.
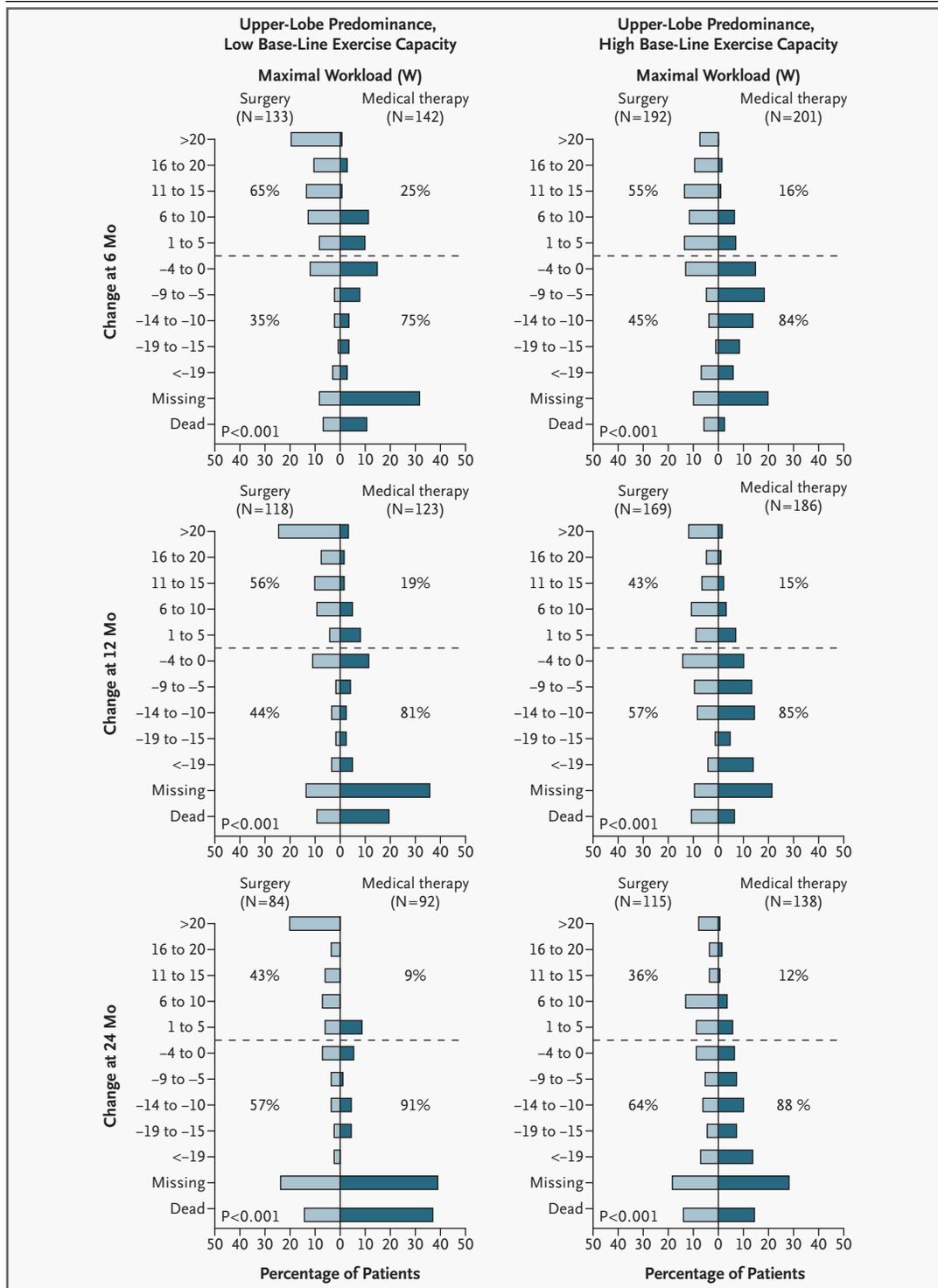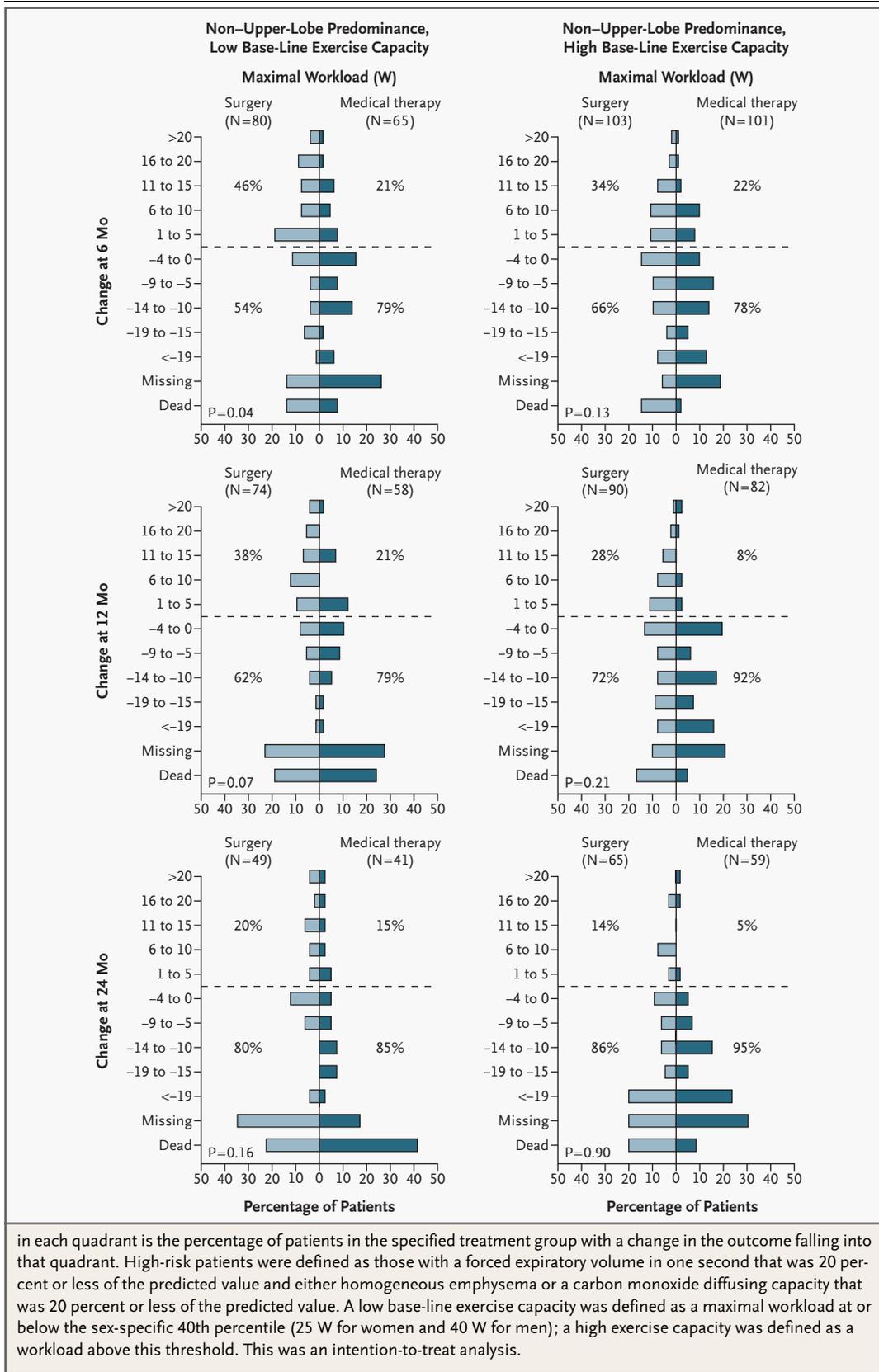
Figure 3. Histograms of Changes from Base Line in Exercise Capacity (Maximal Workload) after 6, 12, and 24 Months of Follow-up in Subgroups of Non–High-Risk Patients.

Base-line measurements were performed after pulmonary rehabilitation. Patients who were too ill to complete the procedure or who declined to complete the procedure but did not explain why were included in the "missing" category. P values were determined by the Wilcoxon rank-sum test. The degree to which the bars are shifted to the upper left of the chart indicates the degree of relative benefit of lung-volume–reduction surgery over medical treatment. The percentage shown

Non–Upper-Lobe Predominance, Low Base-Line Exercise Capacity

Maximal Workload (W)

Non–Upper-Lobe Predominance, High Base-Line Exercise Capacity

Maximal Workload (W)

in each quadrant is the percentage of patients in the specified treatment group with a change in the outcome falling into that quadrant. High-risk patients were defined as those with a forced expiratory volume in one second that was 20 percent or less of the predicted value and either homogeneous emphysema or a carbon monoxide diffusing capacity that was 20 percent or less of the predicted value. A low base-line exercise capacity was defined as a maximal workload at or below the sex-specific 40th percentile (25 W for women and 40 W for men); a high exercise capacity was defined as a workload above this threshold. This was an intention-to-treat analysis.

measures showed a progressive decline from base line in the medical-therapy group; the surgery group had improvements from base line but also had a gradual decline over the course of 24 months (see Supplementary Appendix 7 [available with the full text of this article at http://www.nejm.org]).

PREOPERATIVE PREDICTORS OF OUTCOMES AMONG PATIENTS WITHOUT HIGH RISK

The only individual base-line factors associated with differences in mortality between the treatment groups were the craniocaudal distribution of emphysema (presence or absence of upper-lobe predominance, P for interaction=0.02) and base-line exercise capacity (low or high, P for interaction=0.01). The only individual base-line factor associated with differential improvement in the maximal workload at 24 months was the craniocaudal distribution of emphysema (presence or absence of upper-lobe predominance, P for interaction=0.005). No base-line factor we considered was predictive of differential improvement in the health-related quality of life.

When patients were divided into four subgroups on the basis of combinations of upper-lobe or non–upper-lobe emphysema and low or high exercise capacity at base line, there was strong evidence of differential effects on the risk of death (P for interaction=0.004) and on exercise capacity at 24 months (P for interaction=0.005). Among 290 patients with upper-lobe disease and low exercise capacity, patients in the surgery group had a lower risk of death than patients in the medical-therapy group (risk ratio, 0.47; P=0.005) (Fig. 1D and Table 2), were more likely to have an improvement of more than 10 W in the maximal workload at 24 months (30 percent vs. 0 percent, P<0.001) (Table 3), and were more likely to have an eight-point improvement in the St. George's Respiratory Questionnaire score at 24 months (48 percent vs. 10 percent, P<0.001) (Table 3).

Among the 419 patients with upper-lobe disease and high exercise capacity, mortality was similar, regardless of the treatment-group assignment (risk ratio for death in the surgery group, 0.98; P=0.70) (Fig. 1E and Table 2). However, patients in the surgery group were more likely than those in the medical-therapy group to have improvement of more than 10 W in the maximal workload at 24 months (15 percent vs. 3 percent, P=0.001) (Table 3) and to have an eight-point improvement in the

health-related quality-of-life score at 24 months (41 percent vs. 11 percent, P<0.001) (Table 3).

The 149 patients with non–upper-lobe disease and low exercise capacity had a similar risk of death, regardless of the treatment group (risk ratio for the surgery group, 0.81; P=0.49) (Fig. 1F and Table 2), and had similar chances of improvement of more than 10 W in the maximal workload at 24 months, regardless of the treatment group (12 percent vs. 7 percent, P=0.50) (Table 3), but patients in the surgery group had a greater chance than patients in the medical-therapy group of an eight-point improvement in health-related quality of life at 24 months (37 percent vs. 7 percent, P=0.001) (Table 3).

Among the 220 patients with non–upper-lobe disease and high exercise capacity, patients in the surgery group had a higher risk of death than those in the medical-therapy group (risk ratio, 2.06; P=0.02) (Fig. 1G and Table 2), a similarly low chance of improvement of more than 10 W in the maximal workload at 24 months (3 percent in both groups, P=1.00) (Table 3), and a similar chance of an eight-point improvement in health-related quality of life (15 percent vs. 12 percent, P=0.61) (Table 3).

Changes in the maximal workload are shown in Figure 3 for these four subgroups of patients. For patients with either predominantly upper-lobe emphysema or low exercise capacity after rehabilitation, outcomes favored the surgery group at nearly all time points. Similar patterns of outcomes were seen in the changes in the total daily score for health-related quality of life in these subgroups (see Supplementary Appendix 8 [available with the full text of this article at http://www.nejm.org]). Only in the subgroup with both non–upper-lobe emphysema and high maximal workload at base line did patients in the surgery group not have greater functional and symptomatic improvement than patients in the medical-therapy group.

DISCUSSION

Our study provides reliable estimates of risk and benefit from lung-volume–reduction surgery because of the size of the sample, the use of randomization, the participation of multiple institutions, the use of well-defined measurements, the long-term follow-up, and the low crossover rates. Overall mortality was similar in the surgery and medical-therapy groups, both for all patients and when the previously identified high-risk patients were exclud-

ed. Lung-volume–reduction surgery was associated with a greater chance of improvement in exercise capacity, lung function, quality of life, and dyspnea, but the changes after surgery were highly variable, both among all patients and among those who were not at high risk. After two years, values for the measures of function in survivors in the surgery group had returned nearly to base-line levels, on average, and the values in survivors in the medical-therapy group had continued to deteriorate to levels below base-line values. The functional benefits of lung-volume–reduction surgery came at the price of increased short-term mortality and morbidity.

Because of the broad criteria for inclusion in our study, analysis of prognostic factors allowed us to identify subgroups of patients for whom decisions about lung-volume–reduction surgery are fairly clear-cut. We recognize the pitfalls of subgroup analyses, but we believe that the heterogeneity of the patients and of the outcomes and the considered approach we used make our findings clinically and statistically valid. The subgroup-specific findings were not the result of data mining or the optimization of P values. The candidate prognostic factors we used to identify subgroups were in large part specified in advance on the basis of biologic rationale. The procedures and categorizations for identifying subgroups were carried out under the supervision of the independent data and safety monitoring board.

After high-risk patients had been excluded,[17] four additional clinically meaningful subgroups of patients were identified on the basis of the pattern of emphysema on CT scanning and exercise capacity after rehabilitation. Patients with predominantly upper-lobe emphysema and a low maximal workload after rehabilitation had lower mortality, a greater probability of improvement in exercise capacity, and a greater probability of improvement in symptoms if they underwent lung-volume–reduction surgery than if they received medical therapy alone. In contrast, patients with predominantly non–upper-lobe emphysema and a high maximal workload after rehabilitation had higher mortality if they underwent lung-volume–reduction surgery than if they received medical therapy alone, and they had little chance of functional improvement regardless of the treatment they received. In these subgroups, the risks and benefits of surgery as compared with medical treatment are reasonably clear.

Mortality among the remaining patients was largely independent of the treatment-group assignment. For patients with predominantly upper-lobe disease and a high maximal workload, lung-volume–reduction surgery offered a greater chance for improvement in exercise capacity and symptoms than medical therapy, but only a small percentage of patients had an improvement in exercise capacity of more than 10 W from their base-line level. Patients with non–upper-lobe emphysema and a low maximal workload had only a small chance of improvement in exercise capacity, regardless of the treatment group, but had a greater chance of symptomatic improvement after lung-volume–reduction surgery than with medical therapy.

Upper-lobe emphysema has been associated with short-term improvement in pulmonary function after lung-volume–reduction surgery.[25-27] Our study demonstrates that this pattern of disease is predictive of improved survival in patients who also have low maximal exercise capacity. Upper-lobe predominance, as compared with other patterns, may result in clearer target areas for surgical resection or more accessible areas for excision, or it may indicate that the remaining lung tissue is healthier.

Low exercise capacity was an unanticipated predictor of a survival benefit after lung-volume–reduction surgery. The better survival among patients with low exercise capacity who underwent surgery appears to be due to the very high mortality and marked progressive functional limitation of the patients with low exercise capacity in the medical-therapy group.

Overall, lung-volume–reduction surgery offered no survival benefit. According to our subgroup analysis, it is likely that patients with certain characteristics will have improved survival and function. The survival benefit was limited to patients with predominantly upper-lobe emphysema and a low base-line exercise capacity, but functional benefits were noted in patients with predominantly upper-lobe emphysema and a high base-line exercise capacity and in patients with non–upper-lobe emphysema and a low base-line exercise capacity. Individual outcomes vary widely, but our study provides reliable estimates of outcomes to guide physicians and patients in making decisions about treatment.

APPENDIX

The members of the National Emphysema Treatment Trial Research Group were as follows: *Office of the Chair of the Steering Committee, University of Pennsylvania, Philadelphia:* A.P. Fishman, B.A. Bozzarello, A. Al-Amin. **Clinical Centers:** *Baylor College of Medicine, Houston:* M. Katz, C. Wheeler, E. Baker, P. Barnard, J. Carter, S. Chatziioannou, K. Conejo-Gonzales, J. Haddad, D. Hicks, N. Kleiman, M. Milburn-Barnes, C. Nguyen, M. Reardon, J. Reeves-Viets, S. Sax, A. Sharafkhaneh, C. Young, R. Espada, R. Butanda, K. Dubose, M. Ellisor, P. Fox, K. Hale, E. Hood, A. Jahn, S. Jhingran, K. King, C. Miller, I. Nizami, T. Officer, J. Ricketts, J. Rodarte, R. Teague, K. Williams; *Brigham and Women's Hospital, Boston:* J. Reilly, D. Sugarbaker, C. Fanning, S. Body, S. Duffy, V. Formanek, A. Fuhlbrigge, P. Hartigan, S. Hooper, A. Hunsaker, F. Jacobson, M. Moy, S. Peterson, R. Russell, D. Saunders, S. Swanson; *Cedars–Sinai Medical Center, Los Angeles:* R. McKenna, Z. Mohsenifar, C. Geaga, M. Biring, S. Clark, R. Frantz, P. Julien, M. Lewis, J. Minkoff-Rau, V. Yegyan, M. Joyner; *Cleveland Clinic Foundation, Cleveland:* M. DeCamp, J. Stoller, Y. Meli, J. Apostolakis, D. Atwell, J. Chapman, P. DeVilliers, R. Dweik, E. Kraenzler, R. Lann, N. Kurokawa, S. Marlow, K. McCarthy, P. McCreight, A. Mehta, M. Meziane, O. Minai, P. O'Donovan, M. Steiger, K. White, J. Maurer, C. Hearn, S. Lubell, R. Schilz, T. Durr; *Columbia University, New York, and Long Island Jewish Medical Center, New Hyde Park, N.Y.:* M. Ginsburg, B. Thomashow, P. Jellen, J. Austin, M. Bartels, Y. Berkman, P. Berkoski, F. Brogan, A. Chong, G. DeMercado, A. DiMango, B. Kachulis, A. Khan, B. Mets, M. O'Shea, G. Pearson, J. Pfeffer, L. Rossoff, S. Scharf, M. Shiau, P. Simonelli, K. Stavrolakes, D. Tsang, D. Vilotijevic, C. Yip, M. Mantinaos, M. McKeon; *Duke University Medical Center, Durham, N.C.:* N. MacIntyre, R.D. Davis, J. Howe, R.E. Coleman, R. Crouch, D. Greene, K. Grichnik, D. Harpole, A. Krichman, B. Lawlor, H. McAdams, J. Plankeel, S. Rinaldo-Gallo, J. Smith, M. Stafford-Smith, V. Tapson, M. Steele, J. Norten; *Mayo Foundation, Rochester, Minn.:* J. Utz, C. Deschamps, K. Mieras, M. Abel, M. Allen, D. Andrist, G. Aughenbaugh, S. Bendel, E. Edell, M. Edgar, B. Edwards, B. Elliot, J. Garrett, D. Gillespie, J. Gurney, B. Hammel, K. Hanson, L. Hanson, G. Harms, J. Hart, T. Hartman, R. Hyatt, E. Jensen, N. Jenson, S. Kalra, P. Karsell, D. Midthun, C. Mottram, S. Swensen, A.-M. Sykes, K. Taylor, N. Torres, R. Hubmayr, D. Miller, S. Bartling, K. Bradt; *National Jewish Medical and Research Center, Denver:* B. Make, M. Pomerantz, M. Gilmartin, J. Canterbury, M. Carlos, P. Dibbern, E. Fernandez, L. Geyman, C. Hudson, D. Lynch, J. Newell, R. Quaife, J. Propst, C. Raymond, J. Whalen-Price, K. Winner, M. Zamora, R. Cherniack; *Ohio State University, Columbus:* P. Diaz, P. Ross, T. Bees, H. Awad, J. Drake, C. Emery, M. Gerhardt, M. Kelsey, M. King, D. Rittinger, M. Rittinger; *Saint Louis University, St. Louis:* K. Naunheim, F. Alvarez, J. Osterloh, S. Borosh, W. Chamberlain, S. Frese, A. Hibbit, M.E. Kleinhenz, G. Ruppel, C. Stolar, J. Willey, C. Keller; *Temple University, Philadelphia:* G. Criner, S. Furukawa, A.M. Kuzma, R. Barnette, N. Brister, K. Carney, W. Chatila, F. Cordova, G. D'Alonzo, M. Keresztury, K. Kirsch, C. Kwak, K. Lautensack, M. Lorenzon, U. Martin, P. Rising, S. Schartel, J. Travaline, G. Vance, P. Boiselle, G. O'Brien; *University of California, San Diego, San Diego:* A. Ries, R. Kaplan, C. Ramirez, D. Frankville, P. Friedman, J. Harrell, J. Johnson, D. Kapelanski, D. Kupferberg, C. Larsen, T. Limberg, M. Magliocca, F.J. Papatheofanis, D. Sassi-Dambron, M. Weeks; *University of Maryland at Baltimore, Baltimore, and Johns Hopkins Hospital, Baltimore:* M. Krasna, H. Fessler, I. Moskowitz, T. Gilbert, J. Orens, S. Scharf, D. Shade, S. Siegelman, K. Silver, C. Weir, C. White; *University of Michigan, Ann Arbor:* F. Martinez, M. Iannettoni, C. Meldrum, W. Bria, K. Campbell, P. Christensen, K. Flaherty, S. Gay, P. Gill, P. Kazanjian, E. Kazerooni, V. Knieper, T. Ojo, L. Poole, L. Quint, P. Rysso, T. Sisson, M. True, B. Woodcock, L. Zaremba; *University of Pennsylvania, Philadelphia:* L. Kaiser, J. Hansen-Flaschen, M.L. Geraghty, A. Alavi, T. Alcorn, J. Aronchick, S. Aukberg, B. Benedict, S. Craemer, R. Daniele, J. Edelman, W. Gefter, L. Kotler-Klein, R. Kotloff, D. Lipson, W. Miller, Jr., R. O'Connell, S. Opelman, W. Russell, H. Sheaffer, R. Simcox, S. Snedeker, J. Stone-Wynne, G. Tino, P. Wahl, J. Walter, P. Ward, D. Zisman, J. Mendez, A. Wurster; *University of Pittsburgh, Pittsburgh:* F. Sciurba, J. Luketich, C. Witt, G. Ayres, M. Donahoe, C. Fuhrman, R. Hoffman, J. Lacomis, J. Sexton, W. Slivka, D. Strollo, E. Sullivan, T. Simon, C. Wrona, G. Bauldoff, M. Brown, E. George, R. Keenan, T. Kopp, L. Silfies; *University of Washington, Seattle:* J. Benditt, D. Wood, M. Snyder, K. Anable, N. Battaglia, L. Boitano, A. Bowdle, L. Chan, C. Chwalik, B. Culver, T. Gillespy, D. Godwin, J. Hoffman, A. Ibrahim, D. Lockhart, S. Marglin, K. Martay, P. McDowell, D. Oxorn, L. Roessler, M. Toshima, S. Golden.

Other participants included the following: *Agency for Healthcare Research and Quality, Rockville, Md.:* L. Bosco, Y.-P. Chiang, C. Clancy, H. Handelsman; *Centers for Medicare and Medicaid Services, Baltimore:* S. Sheingold, T. Carino, J. Chin, J. Farrell, K. McVearry, A. Norris, S. Shirey, C. Sikora; *Coordinating Center, Johns Hopkins University, Baltimore:* S. Piantadosi, J. Tonascia, P. Belt, K. Collins, B. Collison, J. Dodge, M. Donithan, V. Edmonds, J. Fuller, J. Harle, R. Jackson, H. Koppelman, S. Lee, C. Levine, H. Livingston, J. Meinert, J. Meyers, D. Nowakowski, K. Owens, S. Qi, M. Smith, B. Simon, P. Smith, A. Sternberg, M. Van Natta, L. Wilson, R. Wise; *Cost-Effectiveness Subcommittee:* R.M. Kaplan, J.S. Schwartz, Y-P. Chiang, M.C. Fahs, A.M. Fendrick, A.J. Moskowitz, D. Pathak, S. Ramsey, S. Sheingold, A.L. Shroyer, J. Wagner, R. Yusen; *Cost-Effectiveness Data Center, Fred Hutchinson Cancer Research Center, Seattle:* S. Ramsey, R. Etzioni, S. Sullivan, D. Wood, T. Schroeder, R. Smith, K. Berry, N. Myers; *CT Scan Image Storage and Analysis Center, University of Iowa, Iowa City:* E. Hoffman, J. Cook-Granroth, A. Delsing, J. Guo, G. McLennan, B. Mullan, C. Piker, J. Reinhardt, J. Sieren, W. Stanford; *Data and Safety Monitoring Board:* J.A. Waldhausen, G. Bernard, D. DeMets, M. Ferguson, E. Hoover, R. Levine, D. Mahler, A.J. McSweeny, J. Wiener-Kronish, O.D. Williams, M. Younes; *Marketing Center, Temple University, Philadelphia:* G. Criner, C. Soltoff; *Project Office, National Heart, Lung, and Blood Institute, Bethesda, Md.:* G. Weinmann, J. Deshler, D. Follmann, J. Kiley, M. Wu.

REFERENCES

1. Cooper JD, Patterson GA, Sundaresan RS, et al. Results of 150 consecutive bilateral lung volume reduction procedures in patients with severe emphysema. J Thorac Cardiovasc Surg 1996;112:1319-30.

2. Criner GJ, Cordova FC, Furukawa S, et al. Prospective randomized trial comparing bilateral lung volume reduction surgery to pulmonary rehabilitation in severe chronic obstructive pulmonary disease. Am J Respir Crit Care Med 1999;160:2018-27.

3. Flaherty KR, Kazerooni EA, Curtis JL, et al. Short-term and long-term outcomes after bilateral lung volume reduction surgery: prediction by quantitative CT. Chest 2001;119:1337-46.

4. Geddes D, Davies M, Koyama H, et al. Effect of lung-volume–reduction surgery in patients with severe emphysema. N Engl J Med 2000;343:239-45.

5. Sciurba FC, Rogers RM, Keenan RJ, et al. Improvement in pulmonary function and elastic recoil after lung-reduction surgery for diffuse emphysema. N Engl J Med 1996; 334:1095-9.

6. Brenner M, Yusen R, McKenna R Jr, et al. Lung volume reduction surgery for emphysema. Chest 1996;110:205-18.

7. Gelb AF, McKenna RJ Jr, Brenner M, Schein MJ, Zamel N, Fischel R. Lung function 4 years after lung volume reduction surgery for emphysema. Chest 1999;116:1608-15.

8. Pompeo E, Marino M, Nofroni I, Matteucci G, Mineo TC. Reduction pneumoplasty versus respiratory rehabilitation in severe emphysema: a randomized study. Ann Thorac Surg 2000;70:948-53.

9. Rationale and design of the National Emphysema Treatment Trial (NETT): a prospective randomized trial of lung volume reduction surgery. J Thorac Cardiovasc Surg 1999;118:518-28.

10. American Thoracic Society. Standardization of spirometry, 1994 update. Am J Respir Crit Care Med 1995;152:1107-36.

11. *Idem.* Single-breath carbon monoxide diffusion capacity (transfer factor): recommendations for a standard technique —

1995 update. Am J Respir Crit Care Med 1995;152:2185-98.

12. Redelmeier DA, Bayoumi AM, Goldstein RS, Guyatt GH. Interpreting small differences in functional status: the Six Minute Walk test in chronic lung disease patients. Am J Respir Crit Care Med 1997;155:1278-82.

13. Steele B. Timed walking tests of exercise capacity in chronic cardiopulmonary illness. J Cardiopulm Rehabil 1996;16:25-33.

14. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation: the St. George's Respiratory Questionnaire. Am Rev Respir Dis 1992;145:1321-7.

15. Kaplan RM, Atkins CJ, Timms R. Validity of a quality of well-being scale as an outcome measure in chronic obstructive pulmonary disease. J Chronic Dis 1984;37:85-95.

16. Eakin EG, Resnikoff PM, Prewitt LM, Ries AL, Kaplan RM. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire: University of California, San Diego. Chest 1998;113:619-24.

17. National Emphysema Treatment Trial Research Group. Patients at high risk of death after lung-volume–reduction surgery. N Engl J Med 2001;345:1075-83.

18. Social Security Administration's Death Master File (full file). Springfield, Va.: National Technical Information Service, 2002. (NTIS order no. SUB-5251INQ.)

19. Social Security Administration Death Master monthly updates file. Springfield, Va.: National Technical Information Service, 2002. (NTIS order no. SUB-5446INQ.)

20. Jones PW. Interpreting thresholds for a clinically significant change in health status in asthma and COPD. Eur Respir J 2002;19:398-404.

21. Stata reference manual, version 7. College Station, Tex.: Stata Press, 2001.

22. Shih JH. Sample size calculation for complex clinical trials with survival endpoints. Control Clin Trials 1995;16:395-407.

23. Agresti A. An introduction to categorical data analysis. New York: John Wiley, 1996.

24. LogXact 4 for Windows: logistic regression software featuring exact methods. Cambridge, Mass.: Cytel Software, 2001.

25. Ingenito EP, Loring SH, Moy ML, et al. Comparison of physiological and radiological screening for lung volume reduction surgery. Am J Respir Crit Care Med 2001;163:1068-73.

26. Wisser W, Senbaklavaci Ö, Özpeker C, et al. Is long-term functional outcome after lung volume reduction surgery predictable? Eur J Cardiothorac Surg 2000;17:666-72.

27. Bloch KE, Georgescu CL, Russi EW, Weder W. Gain and subsequent loss of lung function after lung volume reduction surgery in cases of severe emphysema with different morphologic patterns. J Thorac Cardiovasc Surg 2002;123:845-54.

Copyright © 2003 Massachusetts Medical Society.

RECEIVE IMMEDIATE NOTIFICATION WHEN A *JOURNAL* ARTICLE IS RELEASED EARLY

To be notified when an article is released early on the Web and to receive the table of contents of the *Journal* by e-mail every Wednesday evening, sign up through our Web site at **http://www.nejm.org**